

DOCUMENT RESUME

ED 410 965

IR 056 454

AUTHOR Moen, William E.; McClure, Charles R.
TITLE An Evaluation of the Federal Government's Implementation of
the Government Information Locator Service (GILS). Final
Report.
SPONS AGENCY General Services Administration, Washington, DC. Office of
Information Technology Integration.
PUB DATE 1997-06-30
NOTE 454p.
CONTRACT GS00K96AP2368
PUB TYPE Reports - Evaluative (142)
EDRS PRICE MF01/PC19 Plus Postage.
DESCRIPTORS *Access to Information; Computer Networks; Document
Delivery; *Federal Programs; Information Networks;
*Information Retrieval; Online Systems; Program
Effectiveness; *Program Evaluation; Program Implementation;
Users (Information)
IDENTIFIERS *Government Information Locator Service

ABSTRACT

The Government Information Locator Service (GILS) is an innovative networked-based approach to assist users in locating government information resources. This document reports the results of an evaluation study begun in September 1996 and completed in June 1997. The goal of the study was to understand how: GILS serves various user groups; GILS improves public access to government information; agencies are progressing with their implementations; and how GILS works as a tool for information resources management. A variety of data collection and analysis techniques were used to assess the current status, use, and user satisfaction with the U.S. Federal implementation of GILS. The original vision of a government-wide information locator service has not yet been achieved. Many of the shortcomings of U.S. Federal GILS implementation relate to problems of focus, scope, and administration rather than a fundamental flaw in the architecture and vision of GILS. Some of the issues affecting the success of GILS relate to networked information discovery and retrieval, which represents a large and active research area. GILS, as currently constituted and currently implemented, must be refocused and reengineered to accomplish its original goal as a government-wide information locator service that can improve citizen access to government information. Appendices include: selected policy, management, and operations documents; list of U.S. Federal GILS sites; study methodologies; research instruments; and results of data collection.
(Contains 135 references.) (Author/SWC)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

An Evaluation of the Federal Government's Implementation of the Government Information Locator Service (GILS)

Final Report

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

submitted to

General Services Administration
Office of Information Technology Integration
Washington, DC

Final Deliverable for
Contract No. GS00K96AP2368

submitted by

William E. Moen
University of North Texas
Denton, Texas

and

Charles R. McClure, Distinguished Professor
Manlius, New York

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Charles R. McClure

June 30, 1997

BEST COPY AVAILABLE

DEDICATION

The authors dedicate the work in this study to the memory of Paul Evan Peters, who served as the founding Executive Director of the Coalition for Networked Information. Paul gave the Keynote address for the November, 1996 GILS Conference held at the National Archives and Records Administration, College Park Maryland. That Keynote provided a vision and view for access to networked government information that inspired Conference participants. Paul unexpectedly died later that month.

One of Paul's many interests was U.S. Federal Information Policy and more specifically, improved access to and management of electronic U.S. government information. His interest in and support for the Government Information Locator Service and its ongoing development were well-known. His beliefs, views, and visions regarding access to networked government information have influenced, and will continue to influence many, including these authors.

Acknowledgments

A number of people should be acknowledged for their help and participation in the GILS Evaluation project. The Contracting Officer's Technical Representative (COTR) was Peter Weiss at the Office of Management and Budget. The Co-COTR was Lisa Weber at the National Archives and Records Administration. Individuals that served on the project Advisory Group were Chuck Arnason (Department of Defense), Eliot Christian (U.S. Geological Survey), T. C. Evans (Government Printing Office), Steve Hufford (Environmental Protection Agency), Vicki Ries (Government Printing Office), Lisa Weber, Peter Weiss, and Roxanne Williams (Department of Agriculture). In addition to providing general suggestions regarding the development of the project, they also reviewed and commented on drafts of the final report.

A number of individuals assisted members of the study team in the data collection activities. We especially wish to acknowledge the assistance of those who helped schedule the site visits: Steve Hufford at the Environmental Protection Agency, Vicki Ries and T. C. Evans at the Government Printing Office, Kurt Molholm at the Defense Technical Information Center, and Gladys Myatt and Helen Whatley at the Department of Treasury.

Many other individuals and groups also participated in the study. First there are those who assisted the study team by arranging focus group sessions, interviews, and other data collection activities. In addition, a large number of people participated directly in the study through site visits, interviews, focus groups, completing surveys, and providing other information to the study team. Finally, a number of students assisted in the project by participating in the online user assessment. There simply are too many participants to name individually, but the study team greatly appreciates their contributions.

In addition, a number of individuals were kind enough to review drafts of the final report and provide the principal investigators with suggestions and comments. Again, we greatly appreciate your time and effort.

A nod of thanks goes to other members of the GILS Evaluation study team. J. Timothy Sprehe (Sprehe Information Management Associates), John Carlo Bertot (University of Maryland Baltimore County), Jeff Rubin and Steven K. Wyman (Syracuse University), and June Koelker and Erin Stewart (University of North Texas) all provided a range of assistance with project development, data collection and analysis, report writing, and other aspects of the study. Special thanks goes to Erin Stewart for her tireless work and dedication to this project. Without all their help this project could not have been completed successfully.

The Coalition for Networked Information provided the listserv used by the project Advisory Group throughout the project, and thanks go to Craig Summerhill for his assistance.

We also appreciate the assistance of various staff at the University of North Texas, including the support of Academic Computing Services and the Office of Sponsored Programs. We especially acknowledge the help provided by office and administrative staff in the School of Library and Information Sciences: Paula King, LeAnne Coffey, Greg Straughn, Sharon Willis, and David Zych. And a special thanks to Ginger Boone for computer support, and Kim Kilman and Gloria Shobowale, master's students at the School for their efforts in copyediting, HTML markup, and other project support.

Despite the range of assistance and involvement from others, however, the content of the final report is the responsibility of the principal investigators.

William E. Moen
Charles R. McClure
June, 1997

Executive Summary

The Government Information Locator Service (GILS) is an innovative networked-based approach to assist users in locating government information resources. The U.S. Federal implementation of GILS began in December 1994 with the release of the Office and Management (OMB) Bulletin 95-01. Responsibilities and deadlines prescribed in the OMB Bulletin governed Federal agencies' efforts through a two-year time period (1995-1996). The Bulletin also established a GILS Board "to evaluate the development and operation of the GILS." At its first—and only meeting—in December 1995, the GILS Board approved a recommendation by John Carlin, Archivist of the United States, for an evaluation study of GILS. In his proposal, the Archivist emphasized the importance of understanding how well GILS is meeting user information needs.

This document reports the results of the evaluation study commissioned in response to the GILS Board's request for an assessment of GILS. Five Federal agencies contributed to the funding of the study: Department of Commerce, Department of Defense (DoD), Environmental Protection Agency (EPA), General Services Administration (GSA), and National Archives and Records Administration (NARA). OMB provided the Contracting Officer's Technical Representative. Representatives of eight agencies, including the five sponsoring agencies, served as members of the advisory group to the study and reviewed project plans, findings, and results.

The study began in September 1996, data collection ended in March 1997, and the final report was completed in June 1997. The goal of the study was to understand how:

- GILS serves various user groups
- GILS improves public access to government information
- Agencies are progressing with their implementations
- GILS works as a tool for information resources management.

The principal investigators used a variety of data collection and analysis techniques to assess the current status, use, and user satisfaction with the U.S. Federal implementation of GILS.

Recognizing the complexity of GILS as a networked information service, the investigators considered multiple aspects of GILS, including policy, technology, content, and standards. A primary focus of the study was on users; this focus addressed the charge from the Archivist to examine who is using GILS and how well users' needs are being served, and to identify what modifications are needed to improve service to the public. Data collection activities included site visits to Federal agencies, focus groups with representatives of user communities and stakeholders, policy review, online user assessments of GILS implementations, analysis of Web server transaction logs, and a content analysis of GILS records (see Chapter 2 for the policy review and Chapter 3 for study method). Analysis and synthesis of the data resulted in a series of findings that address the goal of the study and also in a series of recommendations for improving GILS (detailed in Chapter 4).

Agencies' implementation experiences over the two past years have identified issues and challenges that need to be addressed to ensure successful evolution and maturation of GILS. The lessons learned from actual implementation of GILS, which are documented in this report, provide a basis upon which to determine the future shape and character of U.S. Federal GILS implementation.

The investigators conclude that the vision and basic architecture for GILS are still appropriate. The architecture builds on the following components:

- Decentralized deployment of agency-based locators
- Structured and standardized metadata to describe agency information resources
- Z39.50, an American National Standard protocol for information retrieval, for interconnection and interoperable search and retrieval across the agency locators.

The vision of GILS as a service that assists users in locating and accessing publicly-available government information clearly supports important national information policy goals. As originally conceived, a *government-wide* locator

service would result from the separate agency-based GILS.

The investigators conclude, however, that the original vision of a *government-wide* information locator service has not yet been achieved. Rather, there exists a collection of disparate agency GILS that are uneven in their implementation, coverage, and utility. The U.S. GILS implementation has not achieved the vision of a "virtual card catalogue" of government information nor has the majority of agency GILS implementations matured into a coherent and usable government information locator service.

The findings indicate a range of explanations for the current less-than-optimal implementation level of GILS. Many of the shortcomings of U.S. Federal GILS implementation relate to problems of focus, scope, and administration rather than a fundamental flaw in the architecture and vision of GILS. For example, successful GILS implementations were achieved by those agencies that committed sufficient resources, allocated staff, and defined for themselves how GILS could serve their information resources management needs, including the improvement of public access to publicly available agency information. Where an agency has a history of strongly supporting public access to its information resources, GILS tended to be more enthusiastically embraced and perceived as successful than in agencies without such a history. Where top management had endorsed GILS and provided strong support—especially by dedicating staff and capital—GILS tended to be much more successful, at least in its implementation if not in its use.

The study also recognizes that some of the issues affecting the success of GILS relate to networked information discovery and retrieval (e.g., the use of metadata, distributed search and retrieval), which represents a large and active research area. For example, many of the digital library initiatives address, from a research perspective, some of the most complex issues of organization, access, and retrieval of digital information in the global networked environment. Scalable and operational solutions to some of the problems facing networked information locator services have yet to emerge. GILS, as an early innovator, has clarified the nature of some of the problems, and its use of standardized, structured metadata is clearly a contribution.

Findings from a recent evaluation of a Canadian GILS pilot project parallel some of those documented in this report. These parallel findings may indicate some systemic and not clearly understood problems related to networked locators (see Appendix I for the Canadian report). Administrative and organizational commitment from agencies to GILS is a necessary precondition for successful implementation, but solutions to some networked information discovery and retrieval problems may need to emerge before the original vision of GILS is achieved.

The broader government and technological context in which the U.S. Federal GILS implementation occurred also affected agencies' commitment and focus regarding GILS. The U.S. GILS initiative spanned a period of significant technological and agency change, uncertainty, political discord in Washington, opportunity, stress, and excitement for Federal information managers. The last three years have seen more initiatives related to information management and policy than perhaps the last ten years. GILS, given this context, simply was unable to compete for the attention, resources, and commitment from most agency administrators. Three factors in particular—downsizing government, expanding information management legislation and policy issues, and Internet/Web development—should be recognized as affecting the current status of the U.S. Federal GILS initiative.

The investigators organized the findings and recommendations into four primary opportunities:

- Refocus GILS for clarity of purpose and utility
- Improve GILS efficacy in networked information discovery and retrieval
- Resolve GILS relationships with other information handling functions and processes
- Increase GILS awareness.

These opportunities provide a framework for policymakers and implementors to address changes and improvements to the Federal GILS initiative. The table on the following page identifies the findings and recommendations

associated with each opportunity (reported and detailed in Chapter 4).

The first opportunity is where the fundamental decisions and actions for improving GILS should occur and is the primary area for immediate action. This opportunity concerns policy, organizational, and administrative issues that—with appropriate attention and commitment by the GILS Board, OMB, the Chief Information Officers (CIO) Council, and the evaluation study's advisory group—can shape the next phase of GILS evolution. Unlike some of the complex issues related to networked information discovery and retrieval (second opportunity), the policy, organizational, and administrative issues can be resolved. GILS policymakers and implementors can take action to address study findings such as:

- Confused purposes and expectations of what GILS is and should be
- Lack of clear government-wide objectives to guide agencies' implementations
- Expectations for functionality from GILS that were not realistic
- Lack of government-wide coordination, management, and oversight
- Insufficient senior agency management attention and allocation of resources
- Lack of demonstrable benefits to agencies
- A non-workable records management component of GILS.

The implications of these findings bear directly on the users of GILS.

The investigators identified no significant level of user satisfaction with the current U.S. Federal implementation of GILS. Overall, users were confused and disappointed with GILS implementations for a number of reasons, including:

- An inordinately high degree of user sophistication is required to exploit GILS
- Users were interested in and/or expecting to gain access to full-text.
- GILS records were hard to read, contained unnecessary information, and were not linked to the actual source identified

Opportunities, Findings, and Recommendations

Opportunity: Refocus GILS for Clarity of Purpose and Utility

Findings

People Are Confused about GILS Mission, Purposes, and Uses

Expectations for GILS Are Evolving

Government-Wide Administrative Coordination and Policy Oversight Are Lacking

Smaller Agencies Feel Special Burden and Frustration

Agencies' Cultures and Missions Promote Different Commitment to GILS

Intra-Agency Efforts Reflect Different Levels of Enthusiasm for GILS

GILS Benefits Compared to Burdens Are Not Clear

Recommendations

Focus on Public Access to Government Information

Focus Scope of Descriptions On Network-Accessible Information Resources

Identify Responsibilities and Authority for Policy Leadership, Government-Wide Coordination, and Oversight

Implement a Refocused GILS Initiative

Require Agency Reporting on GILS Progress and Reward Agencies That Achieve Stated Objectives

Ensure Ongoing, User-Based Evaluation for Continuous Improvement

Opportunity: Improve GILS Efficacy in Networked Information Discovery and Retrieval (NIDR)

Findings

Web Technology Has Raised Questions about the Role of GILS

GILS is an Agency-Centric, Rather than Government-Wide, Service

GILS Metadata Are Difficult to Capture

Limited Updating and Maintenance of GILS Records

No Clear Agreement on Adequacy of GILS Record Data Elements

Different Types of Resources Represented in GILS Records

User Reaction to GILS Is Not Positive

GILS Record Display Varies Widely and Is Criticized by Users

User Orientation and Instruction is Inadequate

Recommendations

Continuously Evaluate GILS Policies and Standards against Emerging Technologies, Especially the Web

Specify Resource Types And Aggregation Levels

Enforce Consistent Use Of Metadata That Are Empirically Demonstrated to Enhance Networked Information Discovery and Retrieval

Improve Presentation of Metadata

Develop Policy and Procedures for Record Maintenance

Promote Interagency Cooperation and Use of GILS for One-Stop Shopping Functionality

Opportunity: Resolve GILS Relationships with Other Information Handling Functions

Findings

GILS Does Not Support Records Management Activities

GILS Relationship with Agencies' Inventories of Information Resources Is Not Clear

GILS Relationship with FOIA and EFOIA Is Unclear

Recommendations

Uncouple the Refocused GILS—as an Information Discovery and Access Service—from Records Management

Derive GILS Metadata from Other Information Handling Processes

Opportunity: Increase GILS Awareness

Findings

No Program for GILS Promotion and Education Exists

Potential User Communities Lack Familiarity with GILS

GILS Usage Is Limited

Recommendations

Develop and Formalize GILS Promotion, Education, and Training Strategies

- Variance exists in the extent of information contained in GILS records and their display (see Appendix H for two example GILS records that represent this variance).
- The service seemed qualitatively and quantitatively unpredictable and/or uneven.

While a majority of the users reported that they would use GILS to locate government information in the future, there were enough concerns and criticisms from the users to indicate that they consider GILS an unlikely source to help them identify and locate government information.

Knowledge and awareness of GILS in specific important user communities (e.g., government documents librarians) are very limited. If users know of GILS, they make little use of it. When they do use GILS, they find it hard to use at best and inexplicable and frustrating at worst. Even agency staff involved in GILS implementations acknowledged that GILS is "user-unfriendly." Agency staff linked the poor user reception of GILS to difficulties inherent in the search and retrieval system, the lack of full-text information, the limited direct links to the resource when discovered through a GILS record, and deficiencies in marketing GILS.

The current U.S. Federal GILS initiative means different things to different people and has led to inconsistent implementations and a wide range of expectations of GILS. The report makes an overall recommendation that the Federal GILS initiative needs refocusing and alignment with the following vision:

An easy-to-use and coherent government-wide information search service available from one or more service points that enables users to discover, locate, select, and access publicly available government information resources (e.g., agency information systems, specific information dissemination products, and existing locators to those products) through standardized metadata that describe those resources and provide direct links to the described resource (e.g., full-text documents, other online services).

A refocused GILS must clearly articulate the function of a *government-wide* locator service, its scope of coverage, what people can legitimately

expect it to provide, and the benefits it can offer. The purpose of the refocused GILS is to enable users to discover what government information exists and provide users with direct access to that information.

One indication of the more limited scope of a refocused GILS is the investigators' recommendation that GILS and records management should be uncoupled. The current policy identifies GILS as a tool for records management. The study concludes, however, that GILS does not support records management activities. Further, expectations for GILS functionality (e.g., addressing electronic freedom of information requests) that are beyond a primary purpose of assisting users to discover and access government information should be tempered until such functionality can be demonstrated through pilot or prototype implementations.

The refocused GILS effort needs to be clearly distinguished from the early implementation period that was guided by OMB Bulletin 95-10. This demarcation is necessary to 1) acknowledge lessons learned from the early implementation, and 2) acknowledge the frustrations felt by many agencies towards the confused purpose, lack of utility, and limited benefits of many GILS implementations. Government-wide coordination of, identification of realistic objectives for, and education of agencies and users about the refocused GILS are necessary steps in evolving to the next stage of GILS deployment.

The study finds that OMB Bulletin 95-01 was a good first effort to outline a policy context for the development of GILS. Some issues that will require attention in a forthcoming revision to the Bulletin include:

- Clarifying purpose and objectives of GILS
- Divesting records management responsibilities and activities from GILS
- Clarifying Federal leadership for a range of GILS activities
- Recognizing the extent to which agencies can take on GILS responsibilities in a time of budget reductions and increased demands on productivity

- Indicating realistic and tangible benefits that can accrue from GILS
- Integrating GILS into a broader context of agency information systems (including Web sites), information resources management, and general information management missions
- Providing regular oversight and enforcement of GILS policies
- Promoting the development of search and retrieval mechanisms and processes that integrate and coordinate agency components of GILS into a *government-wide* GILS.

To assist government policymakers and implementors, the investigators recommend an initial set of actions to move toward specific solutions and to encourage the success of the refocused GILS across the Federal government. The framework for action includes the following four high priority items:

- Build consensus on the purposes, goals, and scope of the refocused GILS
- Identify who has authority, who is responsible, and where accountability will rest for GILS—as a *government-wide* initiative
- Develop policy goals for GILS and translate them into specific, realistic, and measurable objectives
- Establish a GILS pilot program to identify problems and issues in both policy and implementation arenas.

These four priorities are critical steps to move to the next stage of GILS evolution. Ongoing and continuous evaluation should characterize the refocused GILS effort.

A key first step will be determining who will lead the discussion regarding the future of GILS. The investigators view the evaluation study's advisory group as having responsibilities to review and discuss this evaluation report and then plan the direction for action. The investigators also recommend that the GILS Board, with advice from

the CIO Council and OMB, establish a GILS Transition Task Force to address the four priorities listed above and more specifically, the findings and recommendations in Chapters 4 and 5.

In addition to evaluating and documenting the U.S. Federal GILS implementation and providing decisionmakers with a basis for determining the shape and direction of the next phase of GILS, an additional benefit of this study was the development and refinement of specific techniques for assessing networked information services that agencies can use in ongoing evaluation of their GILS implementations. The intent was to provide policymakers and agency officials with *tools* by which they could deploy a range of assessment techniques and comply with policy such as the Government Performance and Results Act. To date there has been little attention (at least as identified in this study) paid to agency-based performance assessment and the development of performance indicators for GILS efforts. The various instruments developed for this project should be seen as first efforts. Additional research related to these evaluation tools is both necessary and appropriate, and Chapter 5 identifies possible areas for additional effort.

The U.S. Federal government's implementation of GILS has been an ambitious undertaking. Critics may point out limitations and flaws in the current coverage, implementation, and usability of GILS. Equally important, however, is recognizing the progress to date in developing a government information locator service and the commendable efforts by many people who have led and supported GILS implementations. This study recommends that the existing GILS as developed during 1995–1996 be considered as Phase I. The lessons learned from this experience are extensive and can contribute significantly to future efforts to develop a discovery and access service for government information. But GILS, as currently constituted and currently implemented, must be refocused and reengineered to accomplish its original goal as a *government-wide* information locator service that can improve citizen access to government information.

Table of Contents

<i>Acknowledgments</i>	ii
<i>Executive Summary</i>	iii
<i>Chapter 1 Introduction to the Study and Final Report</i>	1
1.0. The Evaluation Study.....	1
1.1. Scope of the Study	3
1.2. Structure of the Report.....	3
1.3. Prelude to the Report	4
1.4. Benefits of the Study.....	6
<i>Chapter 2 Overview and Background on GILS</i>	9
2.0. Introduction.....	9
2.1. A Brief History of GILS	9
2.1.1. Paperwork Reduction, Public Access, and Information Resources Management.....	9
2.1.2. The Emerging Concept of GILS	11
2.1.3. Policy Guidance and Directive in OMB 95-01	13
2.1.4. GILS Implementation.....	15
2.2. The Federal Information Policy Environment for GILS: A Review	16
2.2.1. Overview of Selected Policies	17
2.2.2. Policies Providing Authority for GILS	18
2.2.3. Policies With Intersection and Impact on GILS.....	25
2.2.4. Summary	27
2.3. Selective Review of the Literature.....	28
2.3.1. Literature on Government Information Locators Prior to OMB Bulletin 95-01.....	29
2.3.2. Articles Published after the Issuance of OMB Bulletin 95-01	30
2.3.3. The Post-OMB 95-01 Implementation Period	32
2.3.4. General Themes from the Literature	33
2.4. Policy Findings and Recommendations.....	34
2.5. Lessons from the Policy Analysis and Literature Review	37
<i>Chapter 3 Study Design and Methodology</i>	39
3.0. Introduction.....	39
3.1. Research Methodology Literature Related to Networked Information Services	40
3.2. An Evaluation Framework for GILS	41
3.3. Evaluation Goals and Study Questions.....	43
3.4. The Research Strategy	43

3.5. Multiple Methods and Specific Data Collection Techniques	44
3.5.1. Site Visits	46
3.5.2. Focus Groups	47
3.5.3. Survey	47
3.5.4. Record Content Analysis.....	47
3.5.5. Scripted Online User Assessment	48
3.5.6. Web Server Transaction Log Analysis.....	48
3.5.7. Policy and Literature Review.....	48
3.6. Multiple Dimensions/Multiple Data Events	49
3.6.1. Technology.....	49
3.6.2. Content	49
3.6.3. Users.....	50
3.6.4. Policy	50
3.6.5. Standards and Rules	50
3.7. An Integrated View of GILS.....	51
3.8. Conclusion	51
<i>Chapter 4 Findings and Recommendations</i>	53
4.0. Introduction.....	53
4.1. GILS in the Federal Context.....	56
4.2. Extent of Current GILS Implementation	58
4.3. Opportunity: Refocus GILS for Clarity of Purpose and Utility	63
4.3.1. Finding: People Are Confused about GILS Mission, Purposes, and Uses.....	63
4.3.2. Finding: Expectations for GILS Are Evolving.....	65
4.3.3. Finding: Government-Wide Administrative Coordination and Policy Oversight Are Lacking.....	66
4.3.4. Finding: Smaller Agencies Feel Special Burden and Frustration	68
4.3.5. Finding: Agencies' Cultures and Missions Promote Different Commitment to GILS	68
4.3.6. Finding: Intra-Agency Efforts Reflect Different Levels of Enthusiasm for GILS.....	69
4.3.7. Finding: GILS Benefits Compared to Burdens Are Not Clear	70
4.3.8. Recommendation: Focus on Public Access to Government Information	71
4.3.9. Recommendation: Focus Scope of Descriptions on Network-Accessible Information Resources	72
4.3.10. Recommendation: Identify Responsibilities and Authority for Policy Leadership, Government-Wide Coordination, and Oversight.....	75
4.3.11. Recommendation: Implement a Refocused GILS Initiative	76
4.3.12. Recommendation: Require Agency Reporting on GILS Progress and Reward Agencies That Achieve Stated Objectives	78
4.3.13. Recommendation: Ensure Ongoing, User-Based Evaluation for Continuous Improvement .	79
4.4. Opportunity: Improve GILS Efficacy in Networked Information Discovery and Retrieval (NIDR)	80
4.4.1. Finding: Web Technology Has Raised Questions about the Role of GILS	81
4.4.2. Finding: GILS Is An Agency-Centric, Rather Than Government-Wide, Service	83
4.4.3. Finding: GILS Metadata Are Difficult to Capture	85
4.4.4. Finding: Limited Updating and Maintenance of GILS Records	86
4.4.5. Finding: No Clear Agreement on Adequacy of GILS Record Data Elements.....	86

4.4.6. Finding: Different Types of Resources Represented in GILS Records	87
4.4.7. Finding: User Reaction to GILS Is Not Positive.....	89
4.4.8. Finding: GILS Record Display Varies Widely and Is Criticized by Users	90
4.4.9. Finding: User Orientation and Instruction Is Inadequate	90
4.4.10. Recommendation: Continuously Evaluate GILS Policies and Standards Against Emerging Technologies, Especially the Web	91
4.4.11. Recommendation: Specify Resource Types and Aggregation Levels	92
4.4.12. Recommendation: Enforce Consistent Use of Metadata That Are Empirically Demonstrated to Enhance Networked Information Discovery and Retrieval	94
4.4.13. Recommendation: Improve Presentation of Metadata	95
4.4.14. Recommendation: Develop Policy and Procedures for Record Maintenance.....	96
4.4.15. Recommendation: Promote Interagency Cooperation and Use of GILS for One-Stop Shopping Functionality	97
4.5. Opportunity: Resolve GILS Relationships with Other Information Handling Functions.....	98
4.5.1. Finding: GILS Does Not Support Records Management Activities	98
4.5.2. Finding: GILS Relationship with Agencies' Inventories of Information Resources Is Not Clear.....	100
4.5.3. Finding: GILS Relationship with FOIA and EFOIA Is Unclear.....	101
4.5.4. Recommendation: Uncouple the Refocused GILS-as an Information Discovery and Access Service-from Records Management.....	102
4.5.5. Recommendation: Derive GILS Metadata from Other Information Handling Processes	103
4.6. Opportunity: Increase GILS Awareness	105
4.6.1. Finding: No Program for GILS Promotion and Education Exists.....	105
4.6.2. Finding: Potential User Communities Lack Familiarity with GILS	106
4.6.3. Finding: GILS Usage Is Limited.....	107
4.6.4. Recommendation: Develop and Formalize GILS Promotion, Education, and Training Strategies	110
4.7. Transitioning to the Next Stage of GILS Deployment	111
<i>Chapter 5 Priorities, Next Steps, and Future Research</i>	113
5.0. Introduction.....	113
5.1. Characterizing the Refocused GILS: Recommendations	114
5.1.1. The Refocused GILS Initiative Clearly Articulates the Purposes and Utility of a Government Information Locator Service.....	114
5.1.2. The Refocused GILS Initiative Provides Clear Lines of Authority and Oversight.....	115
5.1.3. The Refocused GILS Initiative Demonstrates Effectiveness and Benefits Through a Pilot Program.....	116
5.1.4. The Refocused GILS Initiative Includes a Program of Government-Wide Education and Promotion.....	117
5.1.5. The Refocused GILS Initiative Emphasizes Continuous Improvement Through Ongoing Evaluation.....	117
5.2. Priorities for Action Towards the Refocused GILS	118
5.3. The Policy Context for the Refocused GILS	119
5.4. Further Research and Evaluation Efforts	119
5.5. Reengineering the GILS Effort.....	122
<i>References</i>	125

Appendices

Selected Policy, Management, and Operations Documents

- Appendix A-1 Office of Management and Budget Bulletin 95-01: Establishment of Government Information Locator Service
- Appendix A-2 Paperwork Reduction Act of 1995 (Section on GILS)
- Appendix A-3 National Archives and Records Administration Bulletin 95-03: Government Information Locator Service
- Appendix A-4 National Archives and Records Administration Bulletin Memorandum: Government Information Locator Service and Privacy Act Notices
- Appendix A-5 Report of the December 6, 1995 Meeting of the Government Information Locator Service (GILS) Board
- Appendix A-6 Statement of Purpose for the OIW/SIG-LA GILS Subgroup

Appendix B List of U.S. Federal GILS Sites as of March 7, 1997

Study Methodologies

- Appendix C-1 Site Visit Methodology
- Appendix C-2 Focus Group Methodology
- Appendix C-3 GILS Conference Survey Methodology
- Appendix C-4 Record Content Analysis Methodology
- Appendix C-5 Scripted Online User Assessment Methodology
- Appendix C-6 Web Server Transaction Log Analysis Methodology

Research Instruments

- Appendix D-1 Site Visit Profile Sheet (Example)
- Appendix D-2 Focus Group Profile Sheet (Example)
- Appendix D-3 GILS Conference Survey Instrument
- Appendix D-4 Record Content Analysis Instrument (Database Fields)
- Appendix D-5 Scripted Online User Assessment Instrument

Results of Data Collection

- Appendix E-1 Summary of GILS Conference Survey Results
- Appendix E-2 Record Content Analysis Findings, Discussion, and Recommendations
- Appendix E-3 Scripted Online User Assessment Findings, Discussion, and Recommendations
- Appendix E-4 Web Server Transaction Log Analysis Results

- Appendix F The Role of GILS Metadata in Networked Information Discovery and Retrieval
- Appendix G Characteristics of Successful Objectives
- Appendix H GILS Record Content and Display Variation
- Appendix I Preliminary Report of the Canadian GILS Subgroup and GILS Pilot Project

List of Tables and Figures

Table 3-1:	Data Collection Techniques	46
Table 4-1:	Opportunities, Findings, and Recommendations	55
Table 4-2:	Record-Source Hosted GILS	59
Table 4-3:	Brokered GILS	60
Table 4-4:	Refocus GILS for Clarity of Purpose and Utility	63
Table 4-5:	Improve GILS Efficacy in Networked Information Discovery and Retrieval	80
Table 4-6:	Resources Described by GILS Records	88
Table 4-7:	Aggregation of Resources Described by GILS	89
Table 4-8:	Resolve GILS Relationships with Other Information Handling Functions	99
Table 4-9:	Increase GILS Awareness	106
Table 4-10:	Familiarity with GILS Documents/Policies	107
Table 4-11:	Use of GILS	108
Table 4-12:	GPO GILS Usage Statistics	109
Table 4-13:	GILS Hits and Accesses as Percentage of Agency HTTP Server.....	110
Figure 2-1:	Policy Perspective on GILS	17
Figure 2-2:	Summary of Selected Policy Instruments Related to GILS	19
Figure 2-3:	Selected Key Issues for GILS Development.....	28
Figure 3-1:	Framework for GILS Evaluation: Perspectives and Dimensions.....	42
Figure 3-2:	Overview of GILS Evaluation Research Methodology	45
Figure 4-1:	Clarifying GILS Purposes and GILS Users	64

Chapter 1

Introduction to the Study and Final Report

At its first meeting in December 1995, the Government Information Locator Service (GILS) Board approved a recommendation by John Carlin, Archivist of the United States, for an evaluation study of GILS. Between September 1996 and March 1997, the investigators conducted extensive data collection and analysis to assess the current status, use, and user satisfaction with the U.S. Federal implementation of GILS. This document is a report of the evaluation study, including findings from the study and recommendations for improving the U.S. Federal GILS initiative.

1.0. THE EVALUATION STUDY

The evaluation of U.S. Federal government's implementation of GILS reported here had as its primary purpose the collection and analysis of information that would lead to an understanding of how:

- GILS serves various user groups
- GILS affects public access to government information
- Agencies are progressing with their implementations
- GILS works as a tool for information resources management.

The Archivist, in his proposal for an evaluation, emphasized the importance of understanding how well "GILS is meeting user information need." He recommended that an evaluation study be conducted that "focuses on who has been using GILS, how well their needs have been served, and what, if any modifications are needed to improve service to the public" (Report of the Initial Meeting of the Government Information Locator Service Board, 12/6/95; See Appendix A-5). The GILS Board established a committee to plan the evaluation.

The General Services Administration contracted with the investigators to conduct the evaluation. Five Federal agencies contributed to the funding of the study: Department of Commerce, Department of Defense (DoD), Environmental Protection Agency (EPA), General Services Administration (GSA), and National Archives and Records Administration

(NARA). The Office of Management and Budget provided support for the study through Peter Weiss as the Contracting Officer's Technical Representative (COTR) Lisa Weber from NARA served as co-COTR. The investigators established a project advisory group, and members of the original evaluation committee served on the advisory group.

The investigators designed and executed a user-based evaluation study that responded to the Archivist's recommendation. The "users" of a government information locator service, however, are not cut from a single cloth, and the investigators accounted for the perspectives of many "user groups," each with a special interest in the U.S. Federal GILS. A user-based approach sensitizes researchers to multiple stakeholders and users with differing needs and expectations, and the effects of these on assessments of programs and services.

The power of a user-based evaluation is its focus on people—their needs, their expectations, and their assessments. User-based evaluations are aligned with user-based design, where the assumption is not *If we build it, they will come* nor that assumption's attendant focus on "systems" and "resources." Instead, user-based design and evaluation focuses on user needs, their behaviors, their requirements, and their assessments of the usability and utility of particular systems and services. While this evaluation study also examined "systems," "resources," and other aspects of the U.S. Federal GILS initiative such as policy and management, users provided a key perspective.

In addition to the extent which GILS implementations by Federal agencies meet the expectations of users, the investigators defined a number of study goals:

- Examine and describe how GILS is serving users in locating and accessing government information.
- Examine and describe agencies' GILS implementation experiences.
- Identify and document success factors and/or barriers affecting agencies' GILS implementations.

- Examine and describe agencies' use of GILS as an information resources management tool.
- Determine if changes to the GILS policies or technical specifications are needed to make it a more useful tool for agency information resources management.
- Provide recommendations and strategies that will assist agencies improve their GILS applications.

The guiding principle for the study was identifying refinements and improvements to the U.S. Federal GILS efforts rather than on measuring strict compliance to policy requirements and technical standards.

The investigators were commissioned to conduct an "Evaluation of the Federal Government's Implementation of the Government Information Locator Service (GILS)" according to the Statement of Work in the General Services Administration's Request for Proposal, KECI-96-006 and based upon the *Technical Proposal* (Moen & McClure, 1996a) submitted in response to the Request for Proposal. The Statement of Work identified specific requirements for the study.

Based on the Statement of Work, the investigators developed the *Work Plan* (Moen & McClure, 1996b) that detailed study activities and time tables. The project advisory group identified above reviewed the *Work Plan*, and the COTR accepted the *Work Plan* as the first deliverable of the study in September 1996. During the study, the advisory group reviewed and provided comments on a progress report (Moen and McClure, 1997), draft preliminary findings and recommendations, and the complete final report.

The investigators have a long history in working with GILS-related activities. Earlier studies included an analysis of locator-related legislation and policy instruments, a survey of existing or planned agency locators, and the design and specification for an agency-based, network-accessible government-wide information locator. Reports from these studies include:

- *The Government Information Locator Service (GILS): Expanding Research and Development on the ANSI/NISO Z39.50 Information Retrieval Standard, Final Report* (Moen & McClure, 1994)
- *Identifying and Describing Federal Information Inventory/Locator Systems: Design for Networked-Based Locators, Volumes I & II* (McClure, Moen & Ryan, 1992)
- *Federal Information Inventory/Locator Systems: From Burden to Benefit* (McClure, et al., 1990).

The investigators brought this knowledge and previous experience with GILS to the current study.

1.1. SCOPE OF THE STUDY

"GILS" as a concept and a vision is broader than any single agency's implementation. As discovered in the study, the term "GILS" means different things to different people. One can use the term to describe a number of things including a generic locator service, the technical specifications for a locator as defined in the GILS Application Profile (see National Institute of Standards and Technology, 1994), or specific implementations and systems providing locator services. The scope of the study became complex because of the range and number of agencies involved in the implementation, the differing views as to what GILS is and should be, and because aspects of GILS, for instance "improving public access," intersect with many other topics and initiatives.

The U.S. Federal implementation of GILS has been directed by policy statements, technical specifications, and implementation guidance. The scope of this evaluation was limited by design and intention to GILS implementations resulting from:

- Office of Management and Budget (OMB) Bulletin No. 95-01, "Establishment of Government Information Locator Service" (1994)
- National Institute of Standards and Technology (NIST) *Federal Information*

- *Processing Standards Publication (FIPS Pub.) No. 192: Application Profile for the Government Information Locator Service (GILS)* (1994)
- National Archives and Records Administration (NARA) *The Government Information Locator Service: Guidelines for the Preparation of GILS Core Entries* (1995).

The investigators use the term "GILS" in the report, unless otherwise specified, with the following meaning:

U.S. Federal implementations of the GILS Application Profile according to specific policy instruments (OMB Bulletin No. 95-01), technical specifications (FIPS Pub. 192), and implementation guidance (NARA's Guidelines for the Preparation of GILS Core Entries).

Terms such as "the Federal GILS initiative," "U.S. Federal GILS," "U.S. implementations of GILS," and "agency GILS," are synonymous with the meaning of "GILS" as defined above.

Any number of other jurisdictions and levels of government are involved in implementations of the GILS Profile. Initiatives at state and international levels often provide innovative approaches for consideration by the Federal government; no doubt this may be true of their GILS efforts. The investigators limited the study to GILS implementations directed by OMB 95-01; other Federal and non-Federal implementations of the GILS Profile were out of scope. Comparative studies of the U.S. Federal GILS and GILS initiatives of states or other national governments would be useful, and the investigators recommend the utility of such comparative studies in Chapter 5.

1.2. STRUCTURE OF THE REPORT

This final report offers findings from the evaluation study and recommendations developed by the investigators to improve the utility of the U.S. Federal GILS initiative. The findings are based on an analysis of the information gained through

various study activities (see Chapter 3). The recommendations address policy, technology, implementation, and other changes to the U.S. Federal implementation of GILS. In addition, the investigators identify near-term and longer-term proposals to move GILS forward.

The chapters following this introduction include background on GILS, a policy and literature review, a description of study methodology, summary of findings and recommendations, discussion of priorities to consider when implementing the recommendations, and identification of areas for further research. Appendices to the report contain relevant GILS documents, detailed descriptions of the study activities, and summary results from various study activities. The appendices are a significant portion of the report and provide sources of data from which the investigators determined findings and formulated recommendations contained in the body of the report.

1.3. PRELUDE TO THE REPORT

The U.S. Federal government's implementation of GILS has been an ambitious undertaking. Critics may point out limitations and flaws in the current coverage, implementation, and usability of GILS. Equally important, however, is recognizing the progress to date in developing a government information locator service and the commendable efforts by many people who have led and supported GILS implementations.

As a mechanism for users to discover, identify, select, and access government information, GILS faced and will continue to face many challenges, including satisfactory resolution of fundamental issues concerning information organization and access that the library and information profession has confronted for many years. In addition, GILS is implemented as a networked information service, and the arena of networked information discovery and retrieval (NIDR) is currently an active research area (e.g. the various digital library initiative). Operational solutions for many NIDR issues do not yet exist.

The Clinton Administration's *National Information Infrastructure: Agenda for Action* intended GILS to be a "virtual card catalogue [sic] that will indicate the availability of government information in whatever form it takes" (Information Infrastructure Task Force, 1993). As happens too often, slogans can both enlighten as well as mask critical issues and challenges. In this case, the reference to the library catalog may obscure the complexity of that mechanism for connecting users with information. As shown later in this report, the complexity of implementing GILS as an agency-based, network-accessible "virtual card catalogue" was significant.

To place GILS development and implementation into perspective, one must recall the past century of library efforts in organizing and providing access to large collections of information. The late 19th century was a vital period for library theoreticians and practitioners who initiated the schemes for information organization and bibliographic control upon which present-day automated and online library information systems are founded. Over the past 100 years, librarians and other information professionals asked fundamental questions about how to connect users with relevant information, especially through the mechanism of the library catalog. They have tried to determine:

- The ways in which users search for information and the access points necessary to support searching
- The information (i.e., metadata) to represent information objects so that they can be discovered, identified, selected, accessed, and used
- The standards necessary to bring consistency to catalogs
- The rules needed to guide the creation of catalog entries
- Mechanisms to link catalogs together effectively.

The answers to these and other information organization and access questions continue to occupy the attention of the library and information science profession. The library's organization and access systems have evolved over the past 100 years because of theoretical and practical knowledge

gained from implementing systems. The experience and lessons learned from efforts to connect users and information have informed each new generation of information organization and access mechanisms.

GILS designers and implementors are addressing long-standing issues related to the organization of and access to government information. Government information may have distinguishing characteristics. Many of the challenges of connecting users to government information, however, are similar to the challenges addressed by librarians and information professionals. GILS designers and implementors, however, have undertaken this initiative in a highly dynamic networked information and technology environment.

The developers of GILS recognized the need for standards to describe and represent government information resources (e.g., the GILS record data elements and structure) and the need for guidelines and rules for the creation of the records (e.g., NARA's *Guidelines*). GILS developers also recognized the need to use evolving information technologies to store, search, and retrieve information (e.g., network technologies and information retrieval protocols). While these aspects of GILS design were—and continue to be—fundamentally appropriate, actual implementation experience can identify problems and raise questions as to the adequacy of even well-considered approaches. Two examples illustrate this point. The structured, standardized GILS records are an important contribution of the GILS initiative, yet in practice the records do not support currently stated goals of GILS for records management. A distributed, decentralized network-accessible locator service is architecturally elegant, but in the actual implementation, U.S. GILS is best characterized as a set of "agency information locators" that taken as a whole do not provide a consistent and coherent view of U.S. government information resources.

The U.S. Federal GILS experience is important in many respects, not the least of which is how agency GILS implementations are highlighting pre-existing conditions (e.g., agency information management practices or the lack thereof) and bringing new problems into finer resolution (e.g., the challenge of

networked information discovery and retrieval, the importance of metadata and the challenges of its capture in a cost-efficient manner). The GILS experience also raises sensitive questions related to decentralization and centralization of information management authority, accountability, and responsibility in the digital age.

GILS was not intended—nor should it try—to provide a single solution to the information organization, access, and management problems of U.S. government information. In the world of information retrieval, many different information systems and services coexist, each with specific purposes and strengths in connecting users with the information they need. The same is true for government information. It is important, however, that GILS does the best job it can according to purposes appropriate for GILS. Defining and specifying what GILS is supposed to do is clearly needed. Most fundamentally, GILS is a "bibliographic instrument" for the networked information environment. GILS can assist users in discovering, identifying, selecting, and accessing U.S. government information. Since GILS is implemented as a networked information service, the early GILS implementation experience has highlighted important issues related to the specifications of bibliographic instruments when used to support networked information discovery and retrieval.

Patrick Wilson, an authoritative voice in the world of information organization, defines bibliographic instruments as having the primary function of listing and describing other writings. Through such instruments, users are able to identify, evaluate, select, and locate information that might be useful to them. In *Two Kinds of Power: An Essay on Bibliographic Control* (1968), Wilson identifies five basic specifications that must be clear—both to designers and users of bibliographic instruments—if they are to have "power" over information:

- The domain of the instrument (in the case of GILS, the domain is government information)
- The principles by which items have been chosen from the domain for inclusion in

- the instrument (e.g., which government information resources will be described in GILS)
- The **unit of analysis** or granularity of the resources that will be described in an entry in the instrument (e.g., what do GILS records describe)
- The **information users can expect** to find in each entry (e.g., what information is consistently given in GILS record)
- The **arrangement and organization** of the instrument (e.g., how to provide coherent views of GILS information?).

While Wilson was writing well before the advent of the networked environment, these five specifications are as pertinent to GILS as to any library catalog. In the coming pages, many of the identified problems with GILS implementations stem, in part, from the lack of understanding and appreciation for these five specifications. To its credit, GILS was an "early adopter" in the arena of networked information discovery and retrieval; concepts and approaches to networked retrieval have been under active development only in the recent years. As noted previously, this arena should be characterized as a research area since many of the problems of distributed search and retrieval have not been solved.

GILS can be seen as a first step in a new regime of the identification and organization of government information resources. If the past 2 years can be seen as an early implementation experiment in this regime, we will be able to look objectively and positively at the lessons learned, identify success factors, and look squarely at the shortcomings and failures. GILS has the potential to address long-standing government information organization and access issues, and the investigators intend this report to be of assistance in reaching that potential. Our recommendations for a refocusing of the GILS effort provides the next evolutionary step in the GILS initiative will continue the work done to date and build upon the experiences and lessons learned for improving public access to government information in the networked environment. We think it is essential, however, that policymakers draw a clear line of demarcation between the early

GILS implementation period (i.e., 1995-1996) and a refocused GILS. One important aspect of such a demarcation is to acknowledge the lessons learned from the early implementation.

1.4. BENEFITS OF THE STUDY

Clearly, the chief product of any evaluation study is the findings and recommendations. In the study reported here, findings and recommendations constitute a major part of the report and will provide directions and strategies for a refocused GILS initiative. Beyond the use and importance of the findings and recommendations, a number of other benefits result from the evaluation study.

First, the study is a statement by the GILS Board, the Office of Management and Budget, and the sponsoring agencies which funded the evaluation that the U.S. Federal GILS implementation is important and deserves a careful review and assessment. Moreover, this statement recognizes that the GILS effort, if it is to develop and improve, needs an external evaluation to guide future decisions and action. In short, the commitment of resources and time to the GILS evaluation is in itself a declaration as to the overall importance of GILS. An easier step to have taken would have been *not* to conduct an evaluation.

Next, the *process* of the evaluation brought together a number of individuals with different perspectives on GILS to exchange information and learn from each other. During the evaluation effort, the investigators were very impressed with the level of interest and involvement in the study by a number of participants, as well as by others who were not study participants but very interested in the outcome of the evaluation. The process of the evaluation brought fresh attention to the GILS effort, raised its visibility within government, and provided a forum to discuss GILS and learn from various implementation experiences.

Third, as policy and evaluation research, the study developed and refined a number of important evaluation techniques that can be used by agencies and others to assess networked information services. Most important, we believe, are: the

progress made on how to conduct Web server log analysis; use of online scripts for user assessments of networked information services such as GILS; and the techniques developed for the GILS record content analysis. Refinements to the methods of focus groups, site visits, surveys, and expert interviews also are important. Appendices to this report describe these techniques in detail.

The depiction of these efforts, reported largely in the appendices, may not do justice to the importance of the techniques undertaken here. Nonetheless, as the investigators discovered at a presentation to the GILS Special Interest Group meeting, April 23, 1997, there is *considerable* interest in these techniques and how to incorporate them as an ongoing part of GILS development and assessment. Participants at the meeting were very interested in applying these techniques to future GILS development at their agencies.

Agencies will need to develop formal measurement and evaluation techniques for their services and systems (such as GILS). This evaluation effort provides a number of useful guidelines and

techniques for agencies developing performance and quality measurement techniques such as required by the *Government Performance Results Act of 1993*. The investigators believe that the assessment techniques and measurements used and tested in this study can be adopted or adapted by agencies for evaluating a variety of networked information systems and services.

Finally, the study provides a formal written assessment of the U.S. Federal GILS effort after roughly 2 years of implementation. As such, it provides a single source of information that all stakeholders can review, discuss, and debate. Whether the evaluation results are taken as benchmarks or beacons, the report provides a foundation for focusing discussions and identifying the work ahead on the beneficiaries of GILS—its users. Regardless of overall agreement with specific findings or recommendations, the report provides a basis for all those interested in the future development of GILS to begin the discussions and move forward with a refocused GILS that will serve agencies and citizen users better.

Chapter 2

Overview and Background on GILS

2.0. INTRODUCTION

An understanding of the current status of U.S. Federal GILS implementation depends in part upon an understanding of the background of the GILS initiative. This chapter provides a brief history of GILS development as well as information about the policy context from which GILS sprang and which continues to affect its existence. Also included in this chapter is a selective review from the professional literature and popular press to indicate the ways in which the U.S. Federal implementation of GILS has been described and interpreted. This chapter, then, provides the overall context from which the investigators began the examination and assessment of GILS.

2.1. A BRIEF HISTORY OF GILS

The concept of a government information locator service emerged from several streams of policy and initiatives within the Federal government dating back to the 1970s. The specifics of the current GILS efforts can be seen as evolving over time, and incorporating along the way the use of networked technologies, changes in information policy directives, and the continuing need of the public to know about and access government information.

2.1.1. Paperwork Reduction, Public Access, and Information Resources Management

The idea for creating some type of locator system for U.S. Federal government information has been in currency for many years. The origins of the current GILS initiative can be traced to information policy efforts, deriving primarily from the work of the Commission on Federal Paperwork (1977) and the *Paperwork Reduction Act* (PRA) of 1980. The 1980 PRA established the Federal Information Locator System (FILS), which was never successfully implemented (Bass & Plocher, 1991). Among the many reasons for its failure was that the statutory formulation of FILS called for a system whose data elements were only based on information collection requests; the scope of the original FILS was quite limited. FILS was a system

for finding and eliminating duplicate Federal information collection requests rather than *locating* information. Oriented toward government information inputs rather than outputs, FILS fell short of functioning as a useful locator system.

Although limited in scope, FILS provided a kernel for the concept of GILS in that it focused on agencies identifying their information resources and making those "inventories" of a limited set of their resources available. The FILS ideas spawned additional ideas and strategies for a government-wide locator system. One approach to locators emerged in the early 1990s with the publication of *Federal Information Inventory/Locator Systems: From Burden to Benefit* (McClure, et al., 1990), a study sponsored by the Office of Management and Budget (OMB). That report called for abandoning FILS and coined the term Government Information/Inventory System (GILS) to describe a new approach that linked inventorying of agency resources and public access.

The 1990 study articulated a specific goal of a government information locator: to enable average citizens to find the government information resources they desired. The idea of a GILS started to receive widespread endorsement both within government circles and within the community of public interest research groups that wanted more and better access to government information. The report included a comprehensive policy review of legislation, regulations, agency guidelines, and other instruments related to government "locator systems" as of 1990.

Another impetus for a locator system was the ongoing efforts by citizens, researchers, librarians, government agencies, and many others to improve access to government information, particularly information in electronic formats. With the increasing amounts of electronic information being generated by the government and the slow pace at which more traditional finding aids for government information kept pace with electronic information, the public needed other mechanisms to assist them in identifying, locating, and accessing government agency information. The new information creation/production environment based on distributed computing and networks also brought

new challenges to traditional models of centralized access to and dissemination of government information (e.g., via the Superintendent of Documents).

Improving public access was a key issue at the 1991 White House Conference on Library and Information Services, where a recommendation was made for the "...federal government to provide comprehensive indexing and abstracting for all public documents to provide easy and equitable access for all individuals" (U.S. National Commission on Libraries and Information Science, 1992, p. 27). Conference attendees considered, but failed to accept, an amendment that urged, "...the federal government to require each agency to maintain an inventory of its publications and urge the federal government to compile and maintain a directory to these agency inventories" (p. 243).

Another outgrowth of the paperwork reduction effort was the development of the information resources management (IRM) concept, which viewed government information resources (both the technology and the data/information) as "assets" that needed to be "managed" as any other agency asset. The key policy statement on managing Federal information resources is the Office of Management and Budget (OMB) Circular A-130, "Management of Federal Information Resources," first issued in 1985 with subsequent revisions in the 1990s (Office of Management and Budget, 1996b). To manage assets adequately, it is necessary to have comprehensive inventories of the assets. Based on the assumption that agencies would inventory their information resources as part of their management of those resources, those inventories could serve as a type of "locator" of information resources.

The link between IRM and enhancing public access to government information became quite clear. Agencies, in the course of implementing IRM policies, would inventory their information resources. Those inventories would be a precondition for adequately managing the resources. Having identified the resources in the inventories, those inventories could be used as a basis for developing finding aids, catalogs, and other locator mechanisms to improve public access to government information. Enhanced public access to

government information would require better and more complete inventories of government information. Information resources inventories could also assist agencies in their records management responsibilities.

In the 1990s, records management also began to emerge as a secondary objective of a GILS. In 1991, the National Historical Publications and Records Commission (1991, p. 13) identified the creation of metadata for managing records as an area for further research:

Research Question 4: How can data dictionaries, information resource directory systems, and other metadata systems be used to support electronic records management and archival requirements?

The report recognized that the metadata information needed to describe and control archival records may be similar to that used by data processing organizations for electronic records management. Descriptive data about agency information resources cast in the form of structured metadata became a centerpiece of the evolving GILS concept.

2.1.2. The Emerging Concept of GILS

Through the early 1990s, the concept of a government-wide information locator service began to take shape. Efforts by Federal agencies as well as two studies by the investigators contributed to the development of the concept.

Among Federal agencies, there was increasing interest in public access issues, in general, as well as interest in the development of some type of a "locator" to government information. One example that began in 1991 was the Interagency Working Group on Public Access, also known as the Solomons Island Group. This group of representatives from a number of Federal agencies met first in May, 1991 (Pesachowitz, 1992), later in November, 1991 (Okay & Williams, 1992), and again in July, 1992 (Phillips & Carroll, 1993). One initiative of the Solomons Island Group was to develop a policy framework for public access to government electronic information. The Working

Group also established subgroups—one of which was "locators and standards"—to further examine policy issues and possible guidelines for locators from an agency perspective. Also during this time period, other Interagency Working Groups such as CENDI explored the development of government wide locator systems. These, and possibly other agency-based efforts, added interest to, credibility about, and an impetus for the development of some type of a Government-wide Information Locator Service (GILS).

The National Archives and Records Administration (NARA), the General Services Administration (GSA), and OMB commissioned a study that resulted in the report, *Identifying and Describing Federal Information/Inventory Locator Systems: Design for Networked-Based Locators* (McClure, Ryan & Moen, 1992). This two volume report made specific recommendations for designing and establishing an agency-based, network-accessible locator system for government information that incorporated a decentralized model for a "virtual" government information locator service. The study recognized the potential of a locator that took advantage of the evolving networked environment. A subsequent discussion of the study, *Design for an Internet-Based Government-Wide Information Locator System* (McClure, Moen, & Ryan, 1992), detailed specific implementation steps for realizing the establishment of government locators in the networked environment.

When the Clinton Administration took office in 1993, a range of government information policy issues quickly took center stage. In its first month, the Administration announced that as part of its technology policy, "We are committed to using new computer and networking technology to make this [government] information more available to the taxpayers who paid for it" (Clinton & Gore, 1993, p. 29). The National Performance Review (NPR) stated that the Administration would, "...require agencies to inventory the federal information they hold, and make it accessible to the public" (Gore, 1993, p. 165). The concept of a government information locator service emerged whereby such a service would be a contribution to the emerging networked infrastructure, both nationally and globally, and would assist the government to do its

job more effectively and efficiently, especially in areas of IRM and public access to information. A report to the Information Infrastructure Task Force in May 1994 crystallized the Administration's thinking on the concept of an information locator system. The report envisioned GILS as a virtual locator, comprised of separate agency-based, network-accessible locators, that used standards for data content and computer communication for interoperable search and retrieval of metadata records (Information Infrastructure Task Force, 1994).

The Clinton Administration's *Agenda for Action* published as part of its National Information Infrastructure (NII) initiative stated (Information Infrastructure Task Force, 1993, p. 3):

The Administration will seek to ensure that Federal agencies, in concert with state and local governments, use the NII to expand the information available to the public, so that the immense reservoir of government information is available to the public, easily and equitably.

These, and other Clinton Administration policy initiatives, incorporated key ideas of GILS into its own information policy strategy.

The movement toward agency inventorying of information required government-wide agreement on a range of standards. These included data content standards for describing information resources as well as standard protocols by which networked systems could communicate, especially for purposes of interoperability of separately implemented agency-based GILS. In Fall 1993, the Public Access Forum Locator Subgroup (of the Solomons Island Group) was developing the content standards for GILS records, and work by the investigators under contract with the U.S. Geologic Survey moved to specify a standards-based technology and data content approach for GILS. The result of the latter work was a report, *The Government Information Locator Service (GILS): Expanding Research and Development on the ANSI/NISO Z39.50 Information Retrieval Standard* (Moen & McClure, 1994). Central to that report was an application profile for GILS (i.e., the GILS Profile) that specified how Z39.50 would be used in

GILS and defined a basic set of data elements comprising a record that would be used to describe agency information resources. The technical specifications for the use of Z39.50 for GILS appeared as *Federal Information Processing Standard Publication (FIPS Pub.) No. 192: Application Profile for the Government Information Locator Service* (National Institute for Standards and Technology, 1994).

ANSI/NISO Z39.50 is the American National Standard that defines a computer protocol for information retrieval (National Information Standards Organization, 1995; see Moen, 1995b for brief introduction to Z39.50). At the time of GILS development, Z39.50 was being routinely implemented by libraries and online information services. GILS, however, was a major new non-library application for Z39.50. In addition, the GILS Profile was one of only two early profiles for use of Z39.50 in a specific application. GILS can be viewed as an early implementor of Z39.50 for non-library applications to achieve interoperability between different computer systems. Ambur discusses a number of issues regarding Z39.50 and interoperability of GILS still to be addressed (1996).

The 1994 report (Moen & McClure, 1994, pp. 16-24) also discussed a number of key policy issues affecting GILS development such as:

- OMB's roles and responsibilities
- GILS and IRM roles and responsibilities
- GILS and record managers' responsibilities
- Technical standards.

To a large degree, these policy issues still affect the overall success of the GILS initiative. Ultimately, the efforts beginning with the Commission on Federal Paperwork (1977), and a range of intervening events, studies, and policy initiatives related to GILS, resulted in the December 1994 OMB Bulletin 95-01, "Establishment of a Government Information Locator Service" (Office of Management and Budget, 1994). The Federal government had now formally mandated a policy on the establishment and operation of GILS. OMB

Bulletin 95-01 is the key policy instrument that currently guides U.S. Federal GILS development and is reprinted as Appendix A-1. In addition, the *Paperwork Reduction Act of 1995* (P.L. 104-13, Sec. 3511) reinforced the executive initiative for GILS through legislative mandate with a section in the Act on the establishment of GILS (see Appendix A-2).

2.1.3. Policy Guidance and Directive in OMB 95-01

FIPS Pub. 192 and The Government Information Locator Service: Guidelines for the Preparation of GILS Core Entries (National Archives and Records Administration, 1995) addressed policy and technical issues related to GILS interoperability and content of GILS records (see Appendix A-3 for NARA Bulletin 95-03). OMB Bulletin 95-01, however, is central and warrants a detailed examination. Simply stated, OMB Bulletin 95-01 directed agencies to implement "agency GILS" that would together comprise the U.S. Federal GILS and result in a government-wide information locator. The Bulletin referenced the *National Information Infrastructure: Agenda for Action* (Information Infrastructure Task Force, 1993) as providing the vision for GILS; *Agenda for Action* called for the establishment of a "virtual card catalogue" [sic] of government information holdings. The Bulletin referenced OMB Circular A-130 for authority in establishing GILS.

The Bulletin's transmittal memo signed by the Director of OMB described the three basic goals for the effort. GILS would:

- Identify public information resources throughout the Federal government
- Describe information available in those resources, and provide assistance in obtaining the information
- Serve as a tool to improve agency electronic records management practices.

The Bulletin articulated agency responsibilities related to GILS, the functions of GILS, and specific requirements for GILS implementation including:

- Identify information resources throughout the Executive Branch
- Describe the information available
- Provide assistance in how to obtain the information
- Improve agencies' abilities to carry out their records management responsibilities and to respond to Freedom of Information Act requests
- Serve to reduce the information collection burden on the public by making existing information more readily available for sharing among agencies.

The Bulletin reflected a vision of GILS as supporting a number of functions (e.g., public access and records management). The applicability of the Bulletin, however, was limited to all departments and agencies in the Executive Branch; independent regulatory commissions and agencies were requested to comply with the Bulletin's mandate. OMB's jurisdiction is limited to these areas of the Federal government, but if GILS does not address Congressional and Judicial information, one can question whether GILS is truly a *government information locator*.

The Bulletin provided definitions of several key GILS concepts including:

- **GILS Core:** "a subset of all GILS locator records which describe information resources maintained by Federal agencies, comply with the GILS core elements defined in Federal Information Processing Standards Publication (FIPS Pub.) 192, and are mutually accessible through interconnected electronic network facilities"
- **Information dissemination product:** "any book, paper, map, machine-readable material, audiovisual production, or other documentary material, regardless of physical form or characteristic, disseminated by an agency to the public" (definition from OMB Circular A-130)
- **Locator:** "an information resource that identifies other information resources, describes the information available in

those resources, and provides assistance in how to obtain the information."

OMB Bulletin 95-01 referenced OMB Circular A-130 not only for the definition of an information dissemination product but also because Circular A-130 directed agencies to maintain inventories of their information resources. The Bulletin stated that such inventories or other finding aids to the resources should be represented in the GILS Core to serve public access and records management goals of GILS. The Bulletin goes on to state, "GILS will become an integral part of the Federal government's overall information management and dissemination infrastructure, and will ultimately facilitate both identification and direct retrieval of government information." Acknowledging the desirability of direct access/retrieval to the information resources described in GILS, the Bulletin expected agencies to do that "to the extent practicable."

OMB Bulletin 95-01 identified a series of GILS implementation activities to be accomplished within specified deadlines (quoted directly from the Bulletin):

- (1) By December 31, 1995, compile an inventory of its 1) automated information systems, 2) Privacy Act systems of records, and 3) locators that together cover all of its information dissemination products. Each such automated information system, Privacy Act system of records, and locator of information dissemination products shall be described by a GILS Core locator record that includes the mandatory GILS Core Elements, and appropriate optional GILS Core Elements as defined in FIPS Pub. 192 and 36 CFR 1228.22(b). Agencies should also supplement the GILS Core Elements with other data elements suitable for specific agency records management and information dissemination needs and objectives. Similar information dissemination products and automated information systems may be identified by a single GILS Core locator record, provided that the locator record clearly identifies the

number and scope of items aggregated. Privacy Act systems of records should, however, be identified individually.

- (2) By December 31, 1995, make its initial GILS Core locator records available online in a form compliant with FIPS Pub. 192 and the related application profile.
- (3) By June 30, 1996, review the information resources identified in the agency inventory of automated information systems and GILS Core locator records for completeness and to determine the extent to which they include Federal records as defined at 44 U.S.C. 3301. For all Federal records covered by the inventory, the agency shall determine whether they are covered by a records disposition schedule authorized by the Archivist of the United States.
- (4) By December 31, 1996, submit to the Archivist a request for disposition authority proposing schedules for unscheduled records in the information resources described in the GILS Core locator records. The agency should also advise the Archivist if it believes any information resource described in the GILS Core locator records has sufficient historical or other value to warrant continued preservation after the information is no longer needed in the agency.

The inventories of agency automated information systems and information dissemination products that are reflected in the GILS Core should serve as the foundation for developing the records schedules proposed by the agency. When an agency needs to retain different categories of records covered by a GILS Core locator record for different periods of time, the agency should supplement the GILS Core locator record by describing each category. Agencies should cite the applicable disposition authority in the GILS Core element for "supplemental

information" for entries that cover records that have been scheduled.

When information dissemination products are part of an on-going series, the agency may submit a proposed records schedule which applies to the entire series. The schedule entry describing such a series may refer to GILS Core locator records to supplement the series description included in the request.

- (5) Continually update its inventory and GILS Core locator records as new information dissemination products and automated information systems are identified.

The Bulletin prescribed these activities and identified deadlines for their accomplishment but was silent on any requirements for agencies to report on their GILS implementations or if they had indeed accomplished the objectives of the directive.

Agencies named by the Bulletin to have special responsibilities for the U.S. Federal GILS initiative included the Department of Commerce (e.g., maintain *FIPS Pub. 192*), NARA (e.g., publish guidance on creating and provide training for using GILS records), and the General Services Administration. The Bulletin also identified "interagency committees" as having an important role in coordinating GILS efforts and developing "interagency topical locators."

The Bulletin also created the Government Information Locator Service Board (the GILS Board), consisting of representatives from a number of government agencies. The GILS Board would "evaluate the development of the GILS," and on an annual basis issue a report that "evaluates and recommends enhancements to GILS to meet user information needs, including factors such as accessibility, ease of use, suitability of descriptive language, as well as the accuracy, consistency, timeliness and completeness of coverage."

In summary, OMB Bulletin 95-01 provided initial policy guidance and direction to agencies in developing their "agency GILS." It focused attention on GILS as functioning in two primary

areas: public access and records management. GILS as a public access device would allow users to identify, locate, and acquire/access information resources of Federal agencies. GILS as a records management device would allow agencies to use GILS records to reduce reporting burdens and facilitate record scheduling.

2.1.4. GILS Implementation

The release of OMB Bulletin 95-01 and the publication of *FIPS Pub. 192* in December 1994 provided the policy and technical specifications for the U.S. Federal GILS implementation. Agencies began developing their implementations in 1995 and continued throughout 1996. Articles noted in Section 2.3 below discuss and describe agency implementation activities throughout this period. Chapter 4 of this report details the extent of agency implementations and identifies a range of issues that are now visible because of this two-year experience.

In December 1995, the GILS Board met for the first—and only time—since the publication of OMB Bulletin 95-01. At that meeting, the Board approved a recommendation for an evaluation of GILS.

NARA hosted a GILS Conference in November 1996 that brought together over 200 people, primarily agency staff but also citizens, academics, and technology vendors. By the time of this Conference, many agencies had had first-hand experiences with GILS. Some came to the Conference with pride in their successful implementations. Others came with concerns about implementation issues, with resentment about having to implement GILS, or with an interest and willingness to implement GILS but confused as to the purpose and definition of GILS. The Conference reflected myriad views of GILS and its future (Baisch, 1997).

GILS has encompassed different meanings as it evolved from conception to implementation. For example, Sally Katzen (1996), the Administrator of the Office of Information and Regulatory Affairs in the Office of Management

and Budget (OMB-ORIA), proposed that GILS can go beyond its original purposes to:

- Become the "killer application" for agencies to use in implementing the provisions of the new *Electronic Freedom of Information Act Amendments of 1996* (EFOIA)
- Become a Global Information Locator Service that binds together the information activities of all governmental entities and their partners worldwide.

Throughout the Conference it was clear that GILS is still in its infancy as far as achieving its *intended* functions as a locator service that promotes and enhances public access to government information and as a records inventorying and scheduling tool to fulfill NARA requirements. People questioned whether GILS should or could support additional functions and expectations such as reducing FOIA requests.

Since the appearance of OMB Bulletin 95-01, however, there has been some controversy as to the role, usefulness, and importance of the GILS initiative. Upon the issuance of OMB Bulletin 95-01, Love wrote that "there is ambiguity over how GILS will work. The system is designed to point to public information resources, but it is unclear how far the system will go in allowing citizens to obtain the documents or data directly" (1994, p. 1). More recently, Henderson identified a range of problems and policy issues and concluded that "only the most minimal expectations were reached in regards to GILS" (1997, p. 1). The current evaluation study was intended to assess agency activities during the past two years and the extent to which GILS is achieving the important policy goals outlined in OMB Bulletin 95-01. Thus, an assessment of the current policy environment related to GILS may provide a useful perspective to help judge the overall effectiveness of the GILS initiative.

2.2. THE FEDERAL INFORMATION POLICY ENVIRONMENT FOR GILS: A REVIEW

The preceding section briefly outlined the development of the concept of GILS leading to OMB Bulletin 95-01 which directed agencies to begin developing agency-based GILS as elements of a government-wide information locator service. Another perspective for understanding GILS is the broader information policy environment in which GILS is being implemented. The purpose of this policy review is to examine selected components of the Federal information policy environment, as of March 1997, as they relate to GILS. More specifically this review has the objectives to:

- Identify the degree to which selected information policy instruments mention, affect, or refer to the GILS initiatives—either implicitly or explicitly
- Compare these policy instruments as to their ambiguity, overlap, contradictions, or gaps as they relate to GILS
- Discuss key issues that arise from the policy review that may require policy attention for the future development of GILS.

The review concentrates on policy instruments developed since 1990 as previous work is available that analyzes Federal policy related to locator systems prior to 1990 (McClure, et al., 1990).

It is important to stress that the policy review presented here is not comprehensive. The policy instruments and initiatives analyzed, in the opinion of the investigators, are key factors that affect the U.S. Federal GILS efforts. More specifically, the policy initiatives discussed in this section do not include those developed by individual agencies. Other sections of this report discuss selected agency policy for GILS development identified as a result of the agency site visits that the investigators conducted during the study.

A well-established fact among information policy analysts is that rather than a single information policy, U.S. Federal information policy is reflected

in a diversity of laws, regulations, directives, and other statements (Heron, McClure & Relyea, 1996). So it was not surprising that subsequent to OMB Bulletin 95-01 in December 1994 establishing GILS, policymakers offered a significant amount of information policy legislation and passed it into law. These laws include the *Government Performance and Results Act of 1993* (U.S. Congress, 1993), the *Paperwork Reduction Act of 1995* (U.S. Congress, 1995), the *Electronic Freedom of Information Act of 1996* (U.S. Congress, 1996a), and the *Information Technology Management Reform Act of 1996* (U.S. Congress, 1996b). Each of these laws, as well as other policy statements such as OMB Circular A-130, either explicitly or implicitly address GILS or GILS functions. The policy environment or context for GILS is dynamic and developed significantly since 1990.

2.2.1. Overview Of Selected Policies

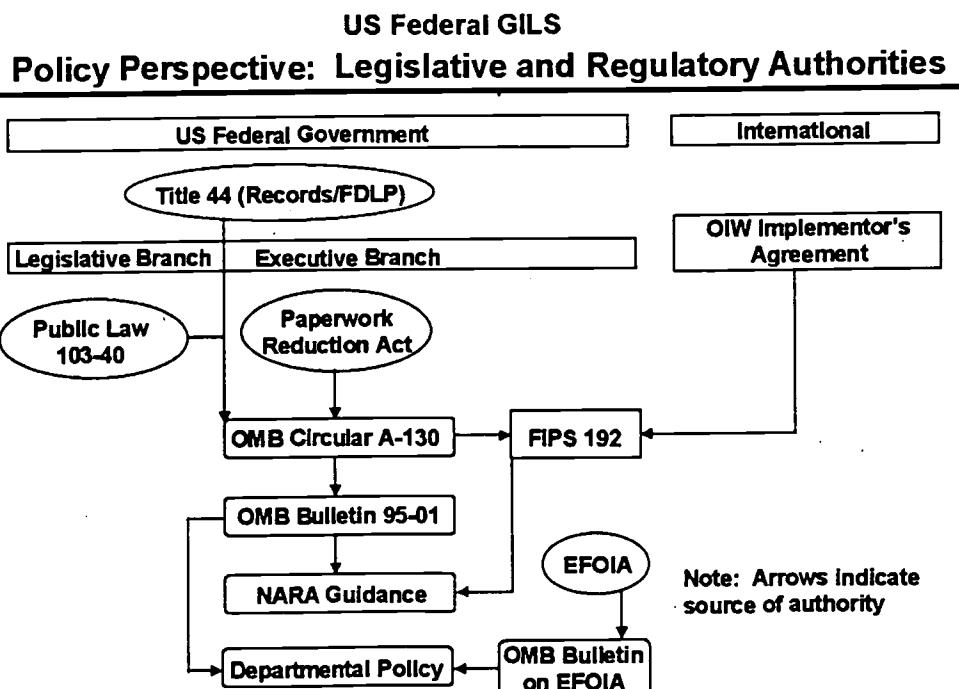
Two areas of policy are of special interest in this review. One area concerns specific policies that provide authority and mandate for the U.S. Federal

GILS initiative. The other area concerns recent legislation and other policy initiatives that can be viewed as intersecting with GILS—either by taking advantage of the existence of agency GILS to support goals of the policy (e.g., EFOIA) or by identifying functions that are GILS-like but do not clearly or explicitly mention GILS.

The discussion above on the historical development of the U.S. Federal GILS initiative identified several efforts since the 1970s that laid a policy and conceptual foundation for GILS. Recent policy statements such as the OMB Bulletin 95-01 came from legislative and regulatory authority. Figure 2-1 presents a policy perspective on U.S. Federal GILS that reflects the linkage and relationships among various policies. (Figure 2-1 is adapted from a graphic developed by the GILS Special Interest Group [GILS SIG] to identify the legislative and regulatory authorities specific to GILS.)

Several specific GILS authorities represented in Figure 2-1 are:

Figure 2-1
Policy Perspective on GILS



- "Establishment of a Government Information Locator Service" (OMB Bulletin 95-01)
- "Management of Federal Information Resources" (OMB Circular A-130)
- *Paperwork Reduction Act of 1995* (P.L. 104-113).

Also associated with these authorities and derived from them are *FIPS Pub. 192* that provided the technical specifications for GILS implementations and NARA's *The Government Information Locator Service: Guidelines for the Preparation of GILS Core Entries*.

Other recent legislation, executive orders, and guidelines can be viewed as intersecting with the U.S. GILS initiative:

- *Government Printing Office Electronic Information Access Enhancement Act of 1993* (Pub. L. 103-40)
- *Electronic Freedom of Information Amendments of 1996* (P.L. 104-231)
- *Information Technology Management Reform Act of 1996* (P.L. 104-106)
- *Government Performance and Results Act of 1993* (P.L. 103-62)
- *Federal Information Technology* (Executive Order 13011, 1996)
- *Guidelines for Agency Use of the World Wide Web* (Office of Management and Budget, 1996a).

Some of these items may not explicitly identify GILS, but they contain important policy implications for GILS. Figure 2-2 provides a side by side analysis summarizing key aspects of the selected information policies identified above.

2.2.2. Policies Providing Authority for GILS

This section briefly summarizes the policy instruments identified in Figure 2-1 that provide authority for GILS.

OMB Bulletin 95-01

OMB Bulletin 95-01 (see Appendix A-1) derives from the authority of Circular A-130, which set forth general policy on information locators in Circular A-130. As discussed in Section 2.1.3., OMB Bulletin 95-01 provided policy goals and direction to agencies regarding U.S. Federal GILS including:

- Scheduling and disposition of records through NARA
- Electronic records management
- Improved agency responses to the Freedom of Information Act (FOIA) requests
- Potential reduction of information collection burden on the public.

The Bulletin states that agencies will be responsible for inventorying and describing holdings and a GILS Board will be established to evaluate the development and operation of GILS and recommend improvements to meet user needs.

OMB Circular A-130

OMB Circular No. A-130 derives from the authority of the *Paperwork Reduction Act* and is the Executive branch implementation of the information policy functions of the Act. While there are no direct references to GILS, Circular A-130 states that agencies are to help the public locate government information by developing retrieval mechanisms for public use (8a(5)(d)(iv)) and establish and maintain inventories of all agency information dissemination products (8a(6)(c)). However, suggestions for these aids are inventories in the form of catalogs and directories, with no specific mention of electronic locators (8a(6)(d)). In its analysis of this policy, Circular A-130 also has some examples of what a locator record might include, such as content, format, means of access, etc. (Appendix IV, analysis of 8a(5)(d)(iv)).

Figure 2-2
Summary of Selected Policy Instruments Related to GILS

Policy	<i>Management of Federal Information Resources OMB Circular A-130 (44 USC 3501 et.seq.)</i>	<i>Establishment of a Government Information Locator Service OMB Bulletin 95-01</i>
Timeline	7/24/94	12/7/94
Direct Reference	No direct language regarding GILS	"GILS will identify information resources throughout the Executive Branch, describe the information available, and provide assistance in how to obtain the information. It will improve agencies' abilities to carry out their records management responsibilities and to respond to Freedom of Information Act requests. It will also serve to reduce the information collection burden on the public by making existing information more readily available for sharing among agencies."
Indirect Reference	<p>"Help the public locate government information maintained by or for the agency" (8a(5)(iv))</p> <p>In Section 8a (6) an information dissemination management system shall at a minimum, "Establish and maintain inventories of all agency information dissemination products"</p> <p>"Develop such other aids to locating agency information dissemination products including catalogs and directories..." (8a(d))</p>	
Responsibility	All Federal agencies	All Federal agencies NARA Archivist – develop GILS core locator elements Secretary of Commerce – designates first Board chair, develops FIPS standard GSA Administrator – make software available to agencies
Oversight	Director of OMB	Government Information Locator Service Board which includes representatives of the OMB Director, the Secretary of Commerce, the Secretary of the Interior, the Archivist of the United States, and the Administrator of General Services. The Public Printer and the Librarian of Congress will be invited to participate as appropriate.
Enforcement	Not really specified, but Director of OMB can grant waivers to agencies	Not specified

Figure 2-2 (cont.)
Summary of Selected Policy Instruments Related to GILS

Policy	<i>Paperwork Reduction Act of 1995 P.L. 104-113 (Amends 44 USC 35)</i>	<i>Government Printing Office Electronic Information Access Enhancement Act of 1993 (Pub. L. 103-40) (Title 44, Sec. 4101 et seq.)</i>
Timeline	5/22/95	6/8/93
Direct Reference	Section 3511. Establishment and operation of Government Information Locator Service “(a)(1) cause to be established and maintained a distributed agency-based electronic Government Information Locator Service...”	No direct language regarding GILS
Indirect Reference		Chapter 41—Access to Federal Electronic Information, Section 4101(a) states that the Superintendent of Documents shall: “(1) maintain an electronic directory of Federal electronic information; (2) provide a system of online access to the Congressional Record, the Federal Register, and, as determined by the Superintendent of Documents, other appropriate publications distributed by the Superintendent of Documents; and (3) operate an electronic storage facility for Federal electronic information to which online access is made available under paragraph (2).”
Responsibility	Section 3511: “(a)(2) require each agency to establish and maintain an agency information locator service as a component of, and to support the establishment and operation of the Service”	All Federal agencies
Oversight	Director of OMB/Administrator of OIRA Interagency Committee – advises Secretary of Commerce on technical standards; will include Archivist of the United States, Administrator of General Services, Public Printer, and the Librarian of Congress	NARA Archivist
Enforcement	Not specified	Not specified

Figure 2-2 (cont.)
Summary of Selected Policy Instruments Related to GILS

Policy	<i>Privacy Act of 1974</i>	<i>Government Performance Results Act of 1993 P.L. 103-62</i>
Timeline	1974	1/5/93
Direct Reference	No direct language regarding GILS.	No direct language regarding GILS.
Indirect Reference	<p>Section 552a(e) states that, "each agency that maintains a system of records shall:"</p> <p>(4) publish in the <i>Federal Register</i> upon establishment or revision a notice of the existence and character of the system of records, which notice shall include –</p> <ul style="list-style-type: none"> (A) the name and location of the system; (B) the categories of individuals on whom records are maintained in the system; (C) the categories of records maintained in the system; (D) each routine use of the records contained in the system, including the categories of users and the purpose of such use; (E) the policies and practices of the agency regarding storage, retrievability, access controls, retention, and disposal of the records; (F) the title and business address of the agency official who is responsible for the system of records; (G) the agency procedures whereby an individual can be notified at his request if the system of records contains a record pertaining to him; (H) the agency procedures whereby an individual can be notified at his request how he can gain access to any record pertaining to him contained in the system of records, and how he can contest its content; and (I) the categories of sources of records in the system. 	<p>No indirect references to information locator systems, but since GILS is a government program, the Act does apply.</p> <p>Section 2(b) states that the Act's purposes are:</p> <ul style="list-style-type: none"> "(3) improve Federal program effectiveness and public accountability by promoting a new focus on results, service quality, and customer satisfaction;" "(4) help Federal managers improve service delivery, by requiring that they plan for meeting program objectives and by providing them with information about program results and service quality;" "(6) improve internal management of the Federal Government."
Responsibility	All Federal agencies	All Federal agencies
Oversight	Congressional committees, Director of OMB	Director of OMB
Enforcement	Director of OMB, though not clearly specified	Congressional budget decisions

Figure 2-2 (cont.)
Summary of Selected Policy Instruments Related to GILS

Policy	<i>Information Technology Management Reform Act of 1996 P.L. 104-106</i>	<i>Federal Information Technology E.O. 13011</i>
Timeline	2/10/96	7/16/96
Direct Reference	No direct language regarding GILS	No direct language regarding GILS
Indirect Reference	<p>Section 5111 (b) highlights the use of information technology</p> <p>“...to improve the productivity, efficiency, and effectiveness of Federal programs, including through the dissemination of public information and the reduction of information collection burdens on the public.”</p> <p>Section 5403 states: “Notwithstanding any other provision of this division, if in designing an information technology system pursuant to this division, the head of an executive agency determines that a purpose of the system is to disseminate information to the public, then the head of such executive agency shall reasonably ensure that an index of information disseminated by such system is included in the directory created pursuant to section 4101 of title 44, United States Code.” [Refers to Government Printing Office’s electronic directory of Federal electronic information.]</p>	<p>No indirect references to information locators.</p> <p>However Section 4(a)(1) states:</p> <p>“creating opportunities for cross-agency cooperation and intergovernmental approaches in using information resources to support common operational areas and to develop and provide shared governmentwide infrastructure services”</p>
Responsibility	Director of OMB/Agency Heads/Chief Information Officers	Agency Heads/Agency Chief Information Officers
Oversight	Director of OMB Secretary of Commerce – Standards and guidelines for computer systems	“seek the views of the Chief Financial Officers Council, Government Information Technology Services Board, Information Technology Resources Board, Federal Procurement Council, industry, academia, and State and local governments on matters of concern to the Council as appropriate.” (Section 3(a)(6))
Enforcement	Director of OMB – budget and appropriations	Office of Management and Budget

Figure 2-2 (cont.)
Summary of Selected Policy Instruments Related to GILS

Policy	<i>Electronic Freedom of Information Amendments of 1996 P.L. 104-231 (Amends 5 USC 552)</i>	<i>OMB Draft Guidelines for Agency Use of the World-Wide Web for Electronic Information Collection, Access and Dissemination, and Management</i>
Timeline	10/2/96	7/16/96
Direct Reference	No direct language regarding GILS.	<p>"Websites shall also include locating aids to any other electronic dissemination and access programs operated by or for the agency. Such programs may include dial-up electronic bulletin boards and third party (intermediary) access services.</p> <p>Full compliance with Government Information Locator Service (GILS) standards will satisfy these locator requirements."</p>
Indirect Reference	<p>"Section 2 (a)(6) Government agencies should use new technology to enhance public access to agency records and information."</p> <p>"Section 2(b)(1) ...ensuring public access to agency records and information"</p> <p>"Section 2(b)(2) improve public access to agency resources and information"</p> <p>"Section 4 (2) a general index of such records [records that have been released to individuals and are likely to have subsequent requests], which shall be made available electronically by December 31, 1999..."</p> <p>Section 11 Directs each agency head to make publicly available upon request, reference material or a guide for requesting records or information from the agency, including: "(1) an index of all major information systems of the agency; (2) a description of major information and record locator systems maintained by the agency; and (3) a handbook for obtaining various types and categories of public information from the agency."</p>	
Responsibility	Agency Heads	All Federal agencies
Oversight	Attorney General – Department of Justice	Internal – Agency established oversight body Suggested members include Chief Information Officer, Public and Congressional Affairs Officers, Records Officer/Manager, Privacy Act Officer, Freedom of Information Act Officer, Security Officer, and appropriate program offices.
Enforcement	U.S. District Court	Existing laws – FOIA, Privacy Act

Paperwork Reduction Act of 1995

In the *Paperwork Reduction Act of 1995* (PRA), Congress wrote into law the establishment and operation of the Government Information Locator Service (P.L. 104-13, Sec. 3511). More importantly, the law clarified that GILS would identify major information systems, holdings, and dissemination products and act as a tool for providing timely, equitable, and useful dissemination of government information to the public. OMB Bulletin 95-01 discussed "information dissemination products," and the GILS records prescribed by the Bulletin were to describe "locators that together cover all of its information dissemination products." The PRA language could be interpreted as prescribing GILS records that would identify specific dissemination products, not simply existing locators of those products. The 1995 PRA also emphasized providing information in a variety of formats, including electronic, and for agencies to make use of available technology.

The PRA also charged the Director of OMB, "in cooperation with the Archivist of the United States, the Administrator of General Services, the Public Printer, and the Librarian of Congress, [to] establish an interagency committee to advise the Secretary of Commerce on the development of technical standards for the Service to ensure compatibility, promote information sharing, and uniform access by the public." OMB 95-01 established the GILS Board with membership to include "representatives of the Director, Office of Management and Budget, the Secretary of Commerce, the Secretary of the Interior, the Archivist of the United States, and the Administrator of General Services" and the Public Printer and the Librarian of Congress were to be "invited to participate as appropriate." Since there is overlapping membership prescribed for these two groups, the investigators queried the study's COTR about the status of the group prescribed by PRA. He responded that the GILS Board "is basically it" (Weiss, 1997). The effectiveness of these two advisory boards—or even if they refer to separate bodies—and the degree to which they have accomplished their stated responsibilities is beyond the scope of the current study.

The evaluation of agency performance, in terms of the requirements of the PRA, is not very detailed or specific, since the agencies only have to present a written report of "steps" taken to improve performance (Sec. 3513, (b)(1)(2)). The further evaluation of performance, that falls on the Director of OMB when reporting to Congress, focuses on describing how collection burdens have been reduced or increased (Sec. 3514, (a)(2)(A)).

National Archives and Records Administration Policies

OMB Bulletin 95-01 directed the NARA to publish guidance and provide training for GILS development. NARA responded with the publication of *The Government Information Locator Service: Guidelines for the Preparation of GILS Core Entries* (National Archives and Records Administration, 1995a), which also outlines how agencies can use GILS to meet their records management responsibilities.

NARA also is committed to the use of GILS for records management. NARA's Strategic Plan for 1997-2007 (National Archives and Records Administration, 1996, p. 11), which addresses how to promote records management, includes the following reference to GILS:

We will emphasize the need for achieving intellectual control, and for scheduling records, by including in our life-cycle management system a means to inventory and schedule records. We will urge federal agencies to use the system so that together we can identify all federal records created by agencies, review their contents, and assure ourselves of not losing essential evidence. We will exploit the Government Information Locator Service to the extent possible for this purpose.

The degree to which this NARA stance, however, has affected policy, and thus, agency records management activities for improved records management is unclear.

2.2.3. Policies with Intersection and Impact on GILS

An aspect of the Federal context in which agencies initiated and pursued GILS implementation is the complexity and depth of information policy issues facing policymakers in the past several years (see also Chapter 4, Section 4.1.) A number of the policies reference locator-like activities (e.g., *Government Printing Office Electronic Information Access Enhancement Act of 1993*), others direct the development of locators (e.g., *Electronic Freedom of Information Act Amendments of 1996*), and still others have little to do with GILS as a locator but intersect at the levels of information technology policy and program accomplishment (e.g., *Information Technology Management Reform Act of 1996* and *Government Performance and Results Act of 1993*). This section summarizes the policy instruments in Figure 2-2 that indirectly impact or intersect with GILS.

Information Technology Management Reform Act of 1996

The *Information Technology Management Reform Act of 1996* (ITMRA) (P.L. 104-106) has the purpose to use information technology (IT) to improve Federal programs. Improvement of programs includes meeting public needs for information and reducing collection burdens (Sec. 5112 (b)). Although ITMRA focuses on acquisitions and information management, it does have bearing on the establishment of GILS. Section 5403 of ITMRA ties information technology systems that disseminate public information to the Government Printing Office (GPO) by requiring that agencies provide information on their systems to the GPO's electronic directory—but “GILS,” *per se*, is not mentioned.

ITMRA emphasizes a fuller integration of IT, “...to promote a coordinated, interoperable, secure, and shared Government-wide infrastructure...” (Section 1 (d)). This is a key aspect to development and functionality of GILS.

As a caveat, a component of this Act is to develop and implement IT standards. Moen and McClure

(1994b) point out that information technology standards should be considered in tandem with information content standards and user needs. They also emphasize that GILS is an example of a standards-based approach to IT development and how, in turn, IT standards support broader information policies (see also Moen, 1994).

ITMRA builds on corporate models by designating a Chief Information Officer (CIO) in all Cabinet and major independent agencies, with primary responsibility for IT management and carrying out PRA functions. CIOs are responsible for monitoring IT performance, including making sure personnel have necessary skills and knowledge to fulfill information resources management duties. ITMRA does not provide any concrete guidance as to how to measure performance, and GILS is not directly mentioned as a tool for improving overall management of information resources or evaluating information resources management practices.

Clearly, GILS can be considered as a tool CIOs may use to carry out their duties. One example of how GILS could be used as a measure of performance is to assess the number of times GILS is used, and for what purposes, via transaction log analyses. (See Appendices C-7 and E-4 for description and results of Web server transaction analysis carried out in this study). This kind of assessment can also be applied to agency fulfillment of the *Government Performance and Results Act of 1993* (GPRA) (P.L. 103-62) requirements for other Federal programs. Yet, ITMRA is conspicuously silent about GILS.

Government Performance and Results Act of 1993

GPRA instituted a requirement for the development of performance measures throughout Executive branch agencies. GPRA seeks to change how agencies assess their programs and services by placing much greater emphasis on what the programs and services are accomplishing, and how well the accomplishments match the programs' purpose and objectives. The advent of GPRA raises the question of whether GILS can and should be viewed within the context of performance measurement. Can Web usage statistics for an

agency's GILS be adapted and considered as measures of agency program performance? It would seem, for example, that the number of "hits," types of hits, sequence of uses, etc. on an agency's GILS would bear some relationship to performance of the agency's information dissemination program goals.

Executive Order 13011, Federal Information Technology

ITMRA was followed by "Federal Information Technology," Executive Order 13011 (1996), that seeks to ameliorate the uncoordinated approach to Federal information resources management by using relevant portions of PRA, ITMRA, and GPRA. The purpose is to improve management and acquisition of information technology in a measurable way, through evaluation of the provision of public service and the degree of agency mission fulfillment. Once again, implications for GILS can be drawn from the Executive Order, but GILS is not explicitly mentioned.

The Executive Order establishes an interagency support structure to, among other things, "minimize unnecessary duplication of effort..." (Section 1 (e)). The inventory capacity of GILS can play a key role as agencies' use of GILS could help them pinpoint similar programs and reduce duplication across information systems.

The establishment of the CIO Council formalizes an interagency support structure. The Council will provide a forum to improve a range of information management practices, including the design, modernization, use, sharing, and performance of information resources. It will also develop recommendations and serve in an advisory capacity to OMB.

Government Printing Office Electronic Information Access Enhancement Act of 1993

The Government Printing Office Electronic Information Access Enhancement Act of 1993 (P. L. 103-40) addressed issues of access to Federal electronic information. Specifically related to

GILS are two provisions of the Act that charge the Superintendent of Documents to:

- Maintain an electronic directory of Federal electronic information
- Operate an electronic storage facility for Federal electronic information to which online access is made available.

This Act became law prior to the establishment of GILS, but clearly agency GILS can serve as a basis for the electronic directory. In fact, with a GILS that meets the goals of OMB 95-01 and PRA, GILS could serve as such a listing of Federal electronic information.

Electronic Freedom of Information Amendments of 1996

The most recent legislation related to GILS is the *Electronic Freedom of Information Act Amendments of 1996* (EFOIA) (P.L. 104-231). This legislation calls for creation and availability of an index of all major information systems of an agency and a "description of major information and record locator systems" maintained by the agency (Section 11). The Act does not refer to GILS, yet what it calls for parallels the intent of GILS. However, it does not create one central point of access/contact for this information, thereby requiring the public to contact individual agencies.

On April 7, 1997 the Director of OMB issued a memorandum providing guidance to address Section 11 of the Act (Office of Management and Budget, 1997a). The memorandum states that agencies can satisfy the requirements listed above for the index and description by establishing a GILS "presence." The lack of precision in the memorandum's language only confuses how GILS and EFOIA can work together, and how, specifically, GILS can assist in handling EFOIA requests.

Privacy Act of 1974

Another policy area to note, that is usually intertwined with FOIA policy, is the *Privacy Act of 1974*. Section 552a(e) states that, "each agency that maintains a system of records shall:"

(4) publish in the *Federal Register* upon establishment or revision a notice of the existence and character of the system of records

The notice is to contain descriptive information about the system of records. The Privacy Act requirements reinforce the role for GILS outlined in OMB Bulletin 95-01 since GILS is intended to increase citizen access to Privacy Act Notices. The current arrangement whereby GPO has mounted a compilation of Privacy Act Notices per the memorandum of agreement between OMB and NARA (see Appendix A-4) to satisfy GILS requirements may need to be specifically addressed by policy guidance in a revised OMB 95-01.

Guidelines for Agency Use of the World Wide Web

OMB has been concerned about the development, management, and operation of Federal agency Web sites and the degree to which they meet existing information policy guidelines. A draft policy statement, *Draft Guidelines for Agency Use of the World-Wide Web for Electronic Information Collection, Access and Dissemination, and Management* (Office of Management and Budget, 1996a) also contained language regarding the role of GILS in agency Web sites. More recently, a draft memorandum from the Administrator of the OMB-OIRA addresses "principles" for Federal agency use of the Web (Office of Management and Budget, 1997b). Formal policy guidance from OMB on this topic remains to be issued. Agency guidelines have been developed by the World Wide Web Federal Web Consortium
<http://www.dtic.mil/staff/cthompson/guidelines/>.

Under "Section V: Additional Points" of the Federal Web Consortium guidelines, the following appears (World Wide Web Federal Consortium, 1996):

GILS is an information processing standard and comprehensive indexing scheme that will identify, describe and help find electronic and non-electronic Federal government information resources. Not only will it point the user to the source of the information; as it evolves, GILS will also provide linkages to assist in its delivery. GILS supplements other agency information dissemination mechanisms and commercial information sources. GILS uses network technology and international standards for information search and retrieval so that information can be retrieved in a variety of ways, and so that GILS users can find other information resources worldwide. Agencies should ensure that a GILS record is created for each agency WWW site. Agencies also should assure that all GILS records which identify WWW-retrievable information dissemination products include linkage to that product. See the DefenseLINK GILS for the DoD implementation at
<http://www.dtic.mil/defenselink/locator/morein.html>.

While such guidelines have no formal authority, one sees the beginning of articulated policy linkage between GILS and the development of agency Web sites.

2.2.4. Summary

This section, as well as the summary offered in Figure 2-2, suggests that there are a number of explicit references to GILS in various policy instruments. There also are a number of *implicit* references to "access to government information," "management of information technology and information resources," and "improving the effectiveness of government operations" in some policy instruments that *could* be inferred to relate to the GILS initiative. The overriding policy goal of enhancing public access to government information is relatively clear throughout many of these laws, executive orders, regulations, etc. It is also clear that these policy instruments attempt some form of intersection of the management of Federal information resources, agency performance, and enhanced access. The resulting policy context of

these instruments, however, is ambiguous since there are instruments that discuss GILS-like functions without referencing GILS, or in the case of PRA and the OMB Bulletin, they can be interpreted as differing in their prescriptions for which resources GILS records should be created.

The 1996 GILS Conference identified a number of issues, many of which clearly have policy implications, that need to be addressed for the

future development of GILS. Figure 2-3 lists a set of issues presented by the investigators at that conference which were largely accepted by the audience as a summary of the key issues affecting the future development of GILS.

The study's data collection activities explored these issues as well as identified other topics and issues. The findings in Chapter 4 describe many of these issues in more detail.

Figure 2-3 Selected Key Issues for GILS Development

- ◆ Agency culture and its attitude toward public access to government information
- ◆ Granularity of the GILS record
- ◆ Making linkages between GILS records and documents records
- ◆ Focus on public access versus records management
- ◆ Content of the GILS record—more, less, different?
- ◆ Need for the automatic generation of GILS records
- ◆ Linking GILS records across government and across servers
- ◆ Policy for better enforcement and oversight of agency GILS activities
- ◆ Clarification between Web Homepage goals and GILS goals—Integration of the two
- ◆ Obtaining user feedback and evaluation of GILS efforts
- ◆ Increasing market demand for GILS efforts
- ◆ Is the current GILS the right product? To what degree should original goals of GILS be revised?
- ◆ Improving buy-in to GILS efforts across the agency and among different agency positions
- ◆ Ensuring top-level support (moral and physical resources) to support GILS efforts
- ◆ Promoting cross-agency interoperability and standards for GILS development

GILS may have the potential to serve multiple information resource management purposes, but is befuddled in some key areas. This is due, in part, to the increased role and importance of information access and the technology to facilitate that process. The fast changing IT environment has increased the difficulty policymakers face in coordinating new information policies with existing ones. An unfortunate side effect of this lack of coordination is a certain degree of confusion that has slowed the agency implementations of GILS. In spite of this, the purpose and open systems structure of GILS is serving as a model for similar services developed by state governments and other countries. While these aspects of GILS are sound, its full realization at the Federal level may require further policy refinement.

2.3. SELECTIVE REVIEW OF THE LITERATURE

Discussions, descriptions, and editorials concerning GILS appear in both the professional and popular literature. The majority of the writings, however, have been descriptive rather than evaluative in nature. This selective review of the literature on GILS focuses on articles and documents about the U.S. Federal GILS effort. A review of the literature on GILS provides an opportunity to identify themes, interpretations, and expectations of GILS. Federal information policy is not covered here since the previous section provided a review of the key policy statements, regulations, and laws related to GILS. The review organizes the literature into three time periods: prior to the release of OMB Bulletin 95-

01, the active implementation phase covered by the Bulletin, and post-December 1996.

2.3.1. Literature on Government Information Locators Prior to OMB Bulletin 95-01

Prior to the publication of OMB Bulletin 95-01 in 1994, articles focused on the need to improve public access through the mechanism of a government-wide information locator. Writers pointed to the problems with existing locators. For example, Bass and Plocher (1991) discussed the aborted attempt of the Federal Information Locator Service (FILS).

Reports from a series of research studies conducted by the investigators at Syracuse University since the early 1990s examined the potential for a government-wide information locator. These research projects produced technical reports (McClure, et al., 1990; McClure, Ryan & Moen, 1992) as well as articles for publication in scholarly and professional journals (McClure, et al., 1991; McClure, Moen & Ryan, 1992). The reports provided thorough background on policy and technical considerations for the development of a government-wide locator. An early design for an agency-based, network-accessible information locator can be found in McClure, Ryan, & Moen (1992).

Christian (1994) offered the first overview in the professional, popular, or scholarly literature of the U.S. Federal GILS concept. Christian situated GILS within a policy context of the 1993 revision of OMB Circular A-130, "Management of Federal Information Resources" and the emerging National Information Infrastructure (NII) efforts of the Clinton Administration. Revisions to OMB Circular A-130 strengthened Federal policy regarding agency responsibilities for information dissemination and encouraged the active management of information by agencies. Christian emphasized the value of public access to government information as indicated by his choice of title, "Helping the Public Find Information: The U.S. Government Information Locator Service." In 1993-94, Christian worked with OMB to refine the concept of GILS, which was documented in a report to the Information Infrastructure Task Force, *The*

Government Information Locator Service (GILS): Report to the Information Infrastructure Task Force (Information Infrastructure Task Force, 1994). (The text of the report is also included in Christian 1996b.)

Sprehe (1994) also positioned GILS within the Clinton Administration's Federal information policy efforts and includes a discussion of the *Freedom of Information Act*, the *Privacy Act*, and the *Paperwork Reduction Act*. Sprehe noted that since the 1980s many agencies had been actively discouraged from a commitment to exploiting their information resources for public benefit. He identified agency public affairs offices as the most likely internal agency consumers for GILS. Sprehe questioned, however, whether a locator system would have enough intrinsic value to produce agencies to cause them to initiate these activities on their own motivation. Sprehe concluded that an imposed requirement (i.e., a GILS mandate) would likely be dropped as soon as external pressure diminished.

Olsen (1994) referenced "GILS as a predecessor to the NII's vision of having desktop 'agents' interact with documents in cyberspace." Olsen quoted Christian (noted above as a principal architect of GILS) as indicating that the tough part of GILS for agencies will be how they decide to represent their information holdings. Olsen's article described a few early agency efforts at information locator services.

In a general overview of Clinton Administration initiatives to make government information more accessible, Thyfault (1994) included GILS as one mechanism among many under consideration at the time. These early plans called for GILS to be available free from kiosks, toll-free phone numbers, electronic bulletin boards, fax, and other off-line media such as floppy disks, CD-ROM, or printed guides.

Overall, the literature reviewed prior to OMB Bulletin 95-01 that addressed the concept of a government-wide information locator service focused heavily on the public access aspect of such a locator service.

2.3.2. Articles Published after the Issuance of OMB Bulletin 95-01

With the release of OMB Bulletin 95-01, writers on GILS could point to actual policy mandate and agency requirements for implementing GILS. OMB Bulletin 95-01 identified specific milestones and deadlines during the period of 1995-1996 for agencies to begin implementing GILS. The National Institute for Standards and Technology (NIST) published *FIPS Pub. 192* in December 1994 which provided the technical specifications to guide agency implementations. Many of the articles in this period provided general background about GILS, described agency implementations activities, or identified how it could be used to improve access to government information. One exception to the descriptive character of most of the writing was a critical review of U.S. Federal GILS implementations by Henderson and McDermott (1995).

Moen (1995) emphasized the value of GILS for the growing geographic information systems (GIS) community. GILS records can describe spatial data, an important set of information resources collected and held by Federal agencies. This article addressed the potential of GILS to provide a means for agencies to manage geospatially referenced information and to assist users in locating spatial data resources held by individual agencies.

One of the more detailed information policy studies published included a reference to GILS within a larger context of public policy and the national information infrastructure (Kalil, 1995). Kalil addressed information policy issues such as privacy, security, and intellectual property within the new digital information environment. He indicated that a Clinton Administration priority was to increase the dissemination of government information. Kalil included GILS in the context of a broader information policy study and highlighted its potential for improving public access to government information. Plocher (1996) linked GILS to larger IRM issues brought about by the passage of the *Paperwork Reduction Act of 1995*.

Henderson and McDermott (1995) surveyed agency GILS implementation efforts as of mid-1995. Their

review included information about the individual agency implementations. They reflected OMB Watch's early dissatisfaction with the general lack of user involvement in the development and implementation of GILS, the lack of uniform or coordinated policy guidance, and the lack of integration between agency Web applications and agency GILS records.

Houser (1995) highlighted the confusion as to the relationship between GILS and the World Wide Web and included reasons why agency staff who build Web pages should implement GILS. He identified GPO as a rival rather than an ally to other Federal agencies and suggested that if agencies don't put up their own GILS records, these records are likely to be housed at GPO, an outcome that Houser implied was undesirable. Houser concluded that agencies benefit from making their information resources accessible through GILS on the Web to promote public access.

Corbin's "Cyberocracy" (1996) reported on the growing scope and importance of agency information available on the Web, and linked the U.S. Federal GILS initiative to broader information access issues stemming from Federal agency use of the Internet. Her statements describing some GILS sites with their impressive search engines were tempered by acknowledging that other GILS sites offered little more than electronic versions of library catalog cards. By distinguishing between effective new options for public access to government information and mere electronic equivalents of limited paper-based information access, Corbin identified a lack of consistency across agencies GILS implementations. This and other articles reflected two emerging themes in the writings on GILS as agency GILS become operational and used by the public. The first was the inconsistency of what resources agencies described in GILS records, and the second issue relates to the relationship between GILS and agency Web applications.

The *Electronic Public Information Newsletter* published articles about GILS in many of its monthly issues in 1995 and 1996. The news articles covered a range of topics including:

- Report on agency compliance with OMB 95-01 (Agencies are generally complying, 1995)
- URL addresses for various agency GILS sites (Information Briefs, 1996a; Information Briefs, 1996c)
- Announcement of plans to survey and evaluate agency GILS (Information Briefs, 1996d)
- Use of GILS by governments outside the U.S. (Information Briefs, 1996b)
- Highlights of the November 1996 GILS conference (GILS conference, 1996)
- Announcement of GILS as a finalist in the Government category for the NII Award (GILS is a finalist, 1996).

These articles provided both the governmental community and the general public with ongoing updates about the GILS activities of Federal agencies.

The November 1996 issue of the *Electronic Public Information Newsletter* summarized issues that surfaced at the 1996 GILS Conference (GILS conference, 1996). These issues included the actual GILS use by the public, the utility of GILS as an Internet information organizing tool, the need to involve end users in GILS design, the need for full access to document level information rather than descriptive records about documents, and the confusion as to the exact kind of information that federal agencies wanted to make accessible to the public.

Two trade newspapers, *Federal Computer Week* and *Government Computer Week* included numerous articles about GILS during the implementation phase. These publications, targeted at technically-oriented government employees and policymakers, included information about GILS efforts in progress at the agencies and other issues related to GILS:

- Hosting of agency GILS records on servers at GPO and NTIS (Jackson, 1996)
- Reference to DTIC's GILS as a "de facto standard for other agencies" (Sikorovsky, 1996b)

- An editorial opinion questioning the use of Z39.50 as the standard for GILS (Temin, 1996)
- Posting of spatial data by USGS for GILS users (Olsen, 1996)
- Lack of coordination between GILS efforts and Web pages (Sprehe, 1996a)
- DOD's requirement to link agency Web pages with GILS records or incur subsequent disconnection of the Web page from the Internet if not linked to GILS (Constance, 1996)
- Support for interoperability between GILS and X.500 (O'Hara, 1996a, O'Hara, 1996b)
- Call for the use of document management systems as an extension of GILS (Varon, 1996)
- A recognition that GILS was proving to be an elusive goal (Power, 1996).

Power (1996) identified the issues of the reliability of data content in GILS records and the importance of public trust in government information sources. He quoted Christian with respect to the last issue as posing the question as to "the electronic equivalent of a royal seal" and "what clues should there be to indicate quality and accuracy of information." Within the context of the U.S. Federal GILS as a method to improve public access to government information, he highlighted the need to address reliability of data (accuracy) and trust as to the source of data (authenticity).

Two NARA employees provided another historical overview on GILS. Adams and Thibodeau (1996) described GILS as a "hallmark of the National Information Infrastructure" and identified three trends which supported the emergence and development of GILS at this point in time:

- Growth in the number of congressional mandates that required Federal agencies to provide public access to specific types of information
- Advances in technology that offered more economical and effective techniques for disseminating electronic information

- Increasing public recognition of the value of government information.

Adams and Thibodeau positioned GILS within a Federal IRM context, which is characterized by the dual functions of access to and management of information. The access component of GILS enhanced public access to government information resources and the management component strengthened agency management of information resources. Their article also discussed the contributions of NARA in establishing descriptive standards for GILS data elements.

The role of the Government Printing Office (GPO) and its relationship to GILS received attention from a number of authors. GPO's actions in support of GILS can be seen in the context of GPO's vision of its future responsibilities in an increasingly electronic publishing environment. Specific articles which linked GILS and GPO are Aldrich (1996), Downing (1996), Farrell, et al. (1996), Gellman (1996), and Sprehe (1996b).

OMB's Bruce McConnell, "New Wine in Old Wineskins, U.S. Government Information in a Networked World" (1996), viewed GILS as a means of locating information in the new networked world. He stated that "information ecology" rather than "information highway" is a more meaningful metaphor and stressed the importance of information being created and sustained in its own niche, connected and interdependent with other information. He supported a distributed responsibility framework for maintaining information in a networked environment and called for creatively managing the evolution of the information ecology.

As the period of intense agency GILS implementations came to a close with the December 31, 1996 deadline prescribed in OMB Bulletin 95-01, Christian, one of the original champions for GILS, published "GILS: What is it? Where's it going?" (Christian, 1996a). Moving beyond the U.S. Federal implementation of GILS, he now situated GILS within the context of a Global Information Infrastructure and highlighted GILS as a means to support decentralized interoperability in

an increasing digital information environment characterized by diversity of sources. Christian presented a vision of GILS as a *Global Information Locator Service* based on design principles including:

- Adoption of open standards
- Support for international use and a diversity of sources
- Implementation within the networked environment
- Recognition of the crucial role of intermediaries
- Access to other locators
- Support of information in different contexts and hierarchies.

Christian identified GILS as a mechanism to provide continuity across different time periods for world data centers as information creation in the future must be able to maintain use of long-term baseline data, using historical, present day, and future data sources interchangeably.

2.3.3. The Post-OMB 95-01 Implementation Period

By the December 31, 1996 deadline identified in OMB Bulletin 95-01, many agencies had completed their initial implementation of GILS. One of the first documents on GILS in 1997 was OMB Watch's second annual report on U.S. Federal GILS implementation (Henderson, 1997). The report recognized that many agencies have either minimally met requirements to implement GILS, and some agencies have done nothing at all. OMB Watch attributed this failure to a lack of specific goals and vision originating from OMB Bulletin 95-01 and to the lack of active involvement of the GILS Board, the oversight body established by the Bulletin. OMB Watch identified specific problems with GILS, including:

- GILS implementations which were limited to Web sites (and not accessible through alternative means)
- Lack of coordination between Web-based full text documents and GILS records

- Limited cataloging of digital resources resulting in less than meaningful electronic access to information described in GILS
- A wide-spread lack of awareness about GILS both among Federal employees and the public.

Ironically, at the time OMB Watch was pointing out problems and failures with the U.S. Federal GILS initiative, Harrel (1997) reported that the Government Information Locator Service was selected as one of the top ten finalists in the 1996 National Information Infrastructure Awards.

Writers on U.S. Federal GILS efforts may now be in a better position to explore and assess the extent to which agency GILS implementations are meeting the goals of OMB 95-01 and whether GILS is satisfying the expectations of various user communities. Actual agency implementations can be examined to see if they can support functions that some expected from GILS. For example, Sprehe (1997) questioned the value of linking GILS records to requirements included in EFOIA. Sprehe distinguished between information publications described by GILS and information contained in government records. EFOIA is intended to provide access to the latter while GILS is intended to provide access to the former. Critiques such as this, as well as assessments by organizations like OMB Watch, and evaluations such as the current study reported here can be a basis for improvements and changes to GILS policy and implementation.

2.3.4. General Themes from the Literature

A number of key themes, issues, and perspectives on GILS emerged from the review of the literature. The investigators conducted ongoing literature review throughout the current evaluation study, and a number of key issues identified in this study are notable by their absence in the literature. The inclusion or absence of these issues in the literature may be indicative of the current strengths and weaknesses of GILS.

Themes identified in the literature include:

- **GILS for Public Access to Government Information:** Most of the GILS literature emphasized the public access aspect of GILS. Many sources cited in the literature review described the potential benefits of GILS for public access to government information. While some writers tie GILS to the management of information resources, the predominant characterization of GILS has been as a tool for improving public access.
- **Confusion about Information Resources Described by GILS Records:** While writers perceive GILS as a means to enhance public access to government information, the GILS literature identified little understanding or agreement on exactly what information was to be described in GILS records.
- **Integration of GILS and the Web:** A number of the articles noted the lack of integration of agency activity in support of Web sites and GILS. It appears that in many agencies, these two activities occurred in parallel with little cross communication, despite the fact that both activities intersect with the electronic access and dissemination of government information. One noticeable exception to the trend was cited in the 1997 OMB Watch report on GILS, which commended EPA for its hotlinks between GILS records and text of documents (Henderson, 1997).

The findings and results of this evaluation study (see Chapter 4) provide an interesting perspective to review what was *not* covered in the GILS literature in the past several years including:

- **Lack of Marketing and Promotional Support for GILS:** As evidenced in the literature, GILS did not have marketing or promotional support from the GILS Board or OMB. From both the professional and popular literature, there were no indications that OMB actively promoted the GILS concept beyond publication of OMB Bulletin 95-01. Within the environment of many Federal agencies, it

may be that marketing or promotional activity of a new service or program is not part of the agencies' cultures. Based on the literature, promotion and marketing of GILS to the agencies and the public was lacking.

- **Minimal Understanding of GILS As a Record Management Tool:** One of the purposes of GILS was to use locator records to enable agencies to better meet record management objectives. The literature contained no critical discussion as to the capability of GILS to support records management.
- **GILS Users:** A key provision of GILS policy was that GILS would enable the public to more effectively identify and access government information. Within the published literature about GILS, little attention is focused on users of GILS records. The absence of articles about such a central provision of the GILS initiative reveals a lack of awareness of the role of users of systems, services, and programs.
- **Management and Coordination of GILS:** What is most telling by its absence in the literature is a lack of coverage of management and coordination issues related to GILS.
- **Evaluation and Assessment of GILS:** The literature was notably silent about how GILS should be evaluated and assessed—either at the agency level or GILS as a government-wide initiative. To date, agencies have concentrated their efforts in becoming GILS compliant. Understanding GILS from an evolutionary perspective, it is possible that once agencies have satisfied the primary requirements to mount GILS records they will then go on to develop programs to evaluate the effectiveness of their efforts in promoting public access to government information. This hopeful outcome is not warranted by information in the GILS literature.

Overall, the literature provided substantial coverage of the U.S. Federal GILS effort, although the identification and critical discussion of issues and

problems with GILS policy and implementation was limited.

2.4. POLICY FINDINGS AND RECOMMENDATIONS

The review of selected policy instruments and literature suggests that there have been considerable discussion and debate about GILS in recent years. Based on the analysis of the policy instruments described in this chapter, there is a clear need to revise the existing OMB Bulletin 95-01. The findings and recommendations offered here are based on the discussion in this chapter. Additional policy findings and recommendations based on the current study's data collection (described in Chapter 3) will be presented in Chapter 4. While specifics for a revision of OMB 95-01 will require careful consideration by OMB-OIRA, the following general areas require attention for such a revision.

Developing a Coherent Policy Environment for GILS

Current information policies and GILS are not well integrated (see Figure 2-2). Steps should be taken to better link GILS into other policy instruments that *implicitly* provide guidelines for access to and management of government information. The implicit references for GILS related activities in ITMRA, E.O. 13011, and Title 44 U.S.C. (Government Printing Office), for example, need to be made *explicit*. It may be appropriate to identify clearly the links to GILS in these and other instruments in a revised OMB Bulletin 95-01.

The coherence of the GILS policy environment suffers from a lack of *explicit* references to GILS, when, in fact, a policy instrument (e.g., ITMRA) deals with topics specifically related to GILS. To some degree a "codification" of guidelines and policies related to GILS (both implicitly and explicitly) may be able to clarify some of the existing ambiguity about GILS and its relationship to other policy instruments.

Purpose and Definition of GILS

OMB Bulletin 95-01 outlined a number of purposes for a GILS and these have been expanded upon over the past 2 years. To some degree, the existing policy environment allows different people to interpret different meanings for the purpose, definition, and content of GILS. Both the public and policymakers have different expectations for what a GILS is, how the GILS is defined, and what a GILS should offer. According to groups like Taxpayer Assets Project and OMB Watch, people expect GILS to provide access to both the records describing government information as well as the full-text of that information.

There has been a pronounced focus on creating GILS records to improve public access, while ignoring the records management functions that GILS was intended to provide for agencies. The detail in OMB Bulletin 95-01 regarding implementation of the records management component of GILS is non-specific and leaves much to individual agency interpretation. Indeed, much of the literature and policy related to GILS stresses the public access aspect of GILS and *not* the records management component. A revised OMB Bulletin 95-01 should specify the purpose and definition of GILS in clear and precise terms.

Role of OMB-OIRA

Inadequate specific guidance exists that clarifies how GILS is to be implemented, the specific nature and content of GILS records, and who has oversight for GILS implementation. On paper, the responsibility for establishing GILS rests with NARA, the Secretary of Commerce, GSA, and each executive agency, with oversight by the Government Information Locator Service Board. In addition, the Government Printing Office, through its online GPO Access program, is one of the primary points where the public encounters GILS.

Specific roles and responsibilities for OMB-OIRA, however, are not made clear in OMB Bulletin 95-01 but should be detailed in a revised policy guideline. Is OMB-OIRA responsible for coordinating these

efforts, for tasking other agencies or groups to be responsible for specific activities, for determining the degree to which others complete their responsibilities? To a large degree there is neither specificity as to OMB-OIRA responsibilities in these areas, nor is it clear as to the enforcement power it has to encourage others to assume GILS responsibilities and implement them effectively.

Role of the GILS Board and Others Groups

OMB Bulletin 95-01 lists responsibilities for the GILS Board. Whether these responsibilities are appropriate or how the Board can promote its recommendations to OMB—or others in the government—is unclear. The administrative relationships among OMB, the GILS Board, various GILS working groups, and other IRM-related groups also is unclear. In addition, if the GILS Board fails to take on its responsibilities as outlined in OMB Bulletin 95-01, who then has oversight responsibility? These relationships, responsibilities, and oversight should be clarified. Specific responsibilities for the GILS Board, the CIO Council, OMB-OIRA, and other agencies/groups (e.g., the GILS SIG) must be better understood if the administration of GILS as a *government-wide* effort is to improve.

Cross-Agency Cooperative Administration of GILS

E.O. 13011 attempted to create cross-agency cooperation, but (disappointingly) does not list the GILS Board or advisory committee as partners from which to "seek views" in terms of oversight. The degree to which agency GILS are administratively linked to other agency GILS—in a federated decentralized fashion as originally conceived—is unclear. Policy guidance for how such linkages should occur and the degree to which a particular agency or an interagency body (e.g., the CIO Council) is to take the "lead" in this area is unclear. It is clear, however, that GILS today is an *agency* information locator service and is not a government-wide locator service. Successful evolution into the next stage of GILS implementation requires policy

guidelines for the overall administrative leadership of GILS.

Enforcement and Oversight of Agency GILS Activities

The visions outlined in the report to the Information Infrastructure Task Force (1994) and in OMB Bulletin 95-01 describe GILS as a decentralized effort by Federal agencies. The lack of explicit supervision and coordination of such decentralized implementations, however, reflects an important policy vacuum. The assumption that people would voluntarily work together to realize the GILS vision appears to be questionable considering the tight budgets, smaller workforces, and larger workloads of the Federal government. Policy guidelines may be needed to clarify enforcement procedures and oversight responsibility to identify those agencies not making good efforts to implement GILS. For example, OMB Bulletin 95-01 has no explicit requirements for agencies to report on their GILS implementation and whether they met the deadlines outlined in the Bulletin. Further, once such agencies are identified, steps that will be taken—and by whom (e.g., OMB-OIRA?)—to obtain the agency's successful participation may also need to be made clear.

Standards for the GILS Record

OMB Bulletin 95-01 indicates that agency GILS will "...contain automated links to underlying databases to permit direct access to information identified in the GILS" (p. 4). This theme is continued in the 1995 PRA where it says that, for information in electronic format, agencies should move towards providing access to the underlying data (P.L. 104-13, Sec. 3506 (d)(1)(B)). But how, exactly will this be done? Who or what has responsibility to determine if the agencies have, in fact, done this or are working on it? And, what assurance is there that the agencies will all develop solutions that are, in fact, interoperable?

The current focus on GILS records within the government has been at the metadata level, where the records serve as pointers to locators of agency

resources, rather than the information itself. The guidance on the "level of detail" or the "specificity" of the locator information, despite the efforts of NARA, is not clear. Policy and direction is needed to clarify the "level of detail" that is appropriate in GILS records, how the standards for records development will evolve, and who or what are responsible for developing and testing such standards. This specificity will enable agencies to develop objectives against which to measure their progress in contributing to the GILS effort.

Relationship Between Agency Web Sites and GILS

There is a definite trend towards providing GILS access via Web sites as evidenced by some Federal agencies, state agencies, and international agencies. It may be time for a standard to be developed, that integrates Z39.50 with the Web, to allow for GILS to be offered via Web servers. The language offered by the World Wide Web Federal Consortium (1996) is a useful first step, but there are a number of policy issues regarding the arrangements between Web sites and GILS that could not be foreseen in December 1994 when 95-01 was developed. Further, the draft Web guidelines proposed by OMB (1996a) may offer a beginning point for integrating Web development with GILS development (Eschenfelder, et al., 1997). Policy guidelines should clarify possible relationships between Web efforts and GILS development. These policies should encourage experimentation and innovation.

User Feedback and Evaluation of GILS Efforts

GILS falls under GPRA as a program for which agencies need to develop performance measures and other assessment techniques. Section 3514 of the PRA ends by stating that any performance evaluation report should be based on "...performance results reported by agencies and shall not increase the collection of information burden on persons outside the Federal Government." This could inhibit the impetus for agencies to solicit user feedback on the usefulness of their GILS. More explicit policy can be developed in a revised OMB Bulletin 95-01 that links the next phase of GILS to GPRA and

encourages performance assessment based on user feedback and assessment.

2.5. LESSONS FROM THE POLICY ANALYSIS AND LITERATURE REVIEW

The policy findings and recommendations offered in this chapter do not constitute a complete set of policy findings and recommendations related to the GILS effort. The investigators developed these findings and recommendations to inform the data collection activities described in Chapter 3. As such, this preliminary list of findings and recommendations shaped the protocol for the site visits, identified questions to be presented at focus group sessions, and clarified issues included in the survey distributed at the November 1996 GILS Conference. Chapter 4 includes additional policy findings and recommendations based on those data collection efforts.

During 1995–1996, after the appearance of OMB Bulletin 95–01, it is interesting to note the limited

attention to GILS—in terms of formal reference—in other key information access and management policies promulgated by the government. No direct reference to GILS in ITMRA, for example, is a lost opportunity to promote GILS into the larger information management community. The lack of mention of GILS as a priority or responsibility for the CIO Council formed in 1996 is also a lost opportunity to promote and extend GILS (Chief Information Officers Working Group, 1996).

To some degree, issues and problems identified with GILS and reported in Chapter 4 have their origin in the policy framework that created GILS, as discussed in this chapter. To the defense of the creators of OMB Bulletin 95–01, the GILS effort was an experiment for which there was limited knowledge about GILS, its creation, and implementation. The actual implementation experiences by agencies in the past 2 years have made a range of GILS issues visible. The current study has identified a number of these issues—reported in this chapter and in Chapter 4—that will require policy revisions if the U.S. Federal implementation of GILS is to be successful.

Chapter 3

Study Design and Methodology

3.0. INTRODUCTION

The evaluation of complex networked-based information services presents unusual challenges to researchers. Evaluations that examine a single dimension or aspect of such information services are likely to be limited in their utility. For this evaluation of U.S. Federal implementations of GILS, the investigators designed and developed a multi-method research approach appropriate for the multi-faceted nature of GILS. The investigators documented the proposed design of the study in the *Technical Proposal* (Moen & McClure, 1996a), which they submitted and had accepted by the contracting agency. Upon award of the contract, the investigators completed a *Work Plan* (Moen & McClure, 1996b) that detailed the research strategy, methodological approach, data collection and analysis activities, and other considerations in carrying out a rigorous assessment of GILS. The project advisory group reviewed and approved the *Work Plan*.

This chapter discusses the design of the study and the multi-method approach used in the evaluation. The chapter also reports on the extent of data collection, including numbers of activities accomplished and participants involved. Appendices C-1 through C-6 provide additional information on the study and contain detailed summaries of each data collection and analysis activity.

The architecture of GILS includes metadata (or pointers) that describe a range of government information resources (electronic as well as non-electronic), human intermediaries, technical standards, government-wide and agency policy, users, and various information technologies. GILS, as a networked information service, reflects a complexity resulting from the interaction of a number of dimensions including policy, content, users, technology, and standards. Evaluation methodology for complex, networked-based information resources is emerging due in part to the ARPA/NASA Digital Library Initiatives (see Allerton Institute, 1995, "How We Do User-Centered Design and Evaluation of Digital Libraries: A Methodological Forum"). The need to develop appropriate tools and methods for

evaluation and assessment of networked services is critical, especially those that incorporate a user-based perspective (McClure, 1994).

For the current evaluation study, the investigators anticipated that the GILS implementation process would differ from agency to agency. Each agency has its own type and quantity of resources to be described in GILS. Additionally, each agency has its own technological infrastructure, individual administrative expertise, and financial resources to implement such a service. These factors, along with the agency's culture, affected each agency's readiness to implement GILS.

Given the multi-faceted characteristic of GILS—policy, technology, standards, content, and users—the investigators crafted an evaluation research approach appropriate to the complex phenomenon of GILS and to the purposes and goals of the evaluation. A review of research methodology literature in the area of networked information services aided the study team in the design of the evaluation methodology.

3.1. RESEARCH METHODOLOGY LITERATURE RELATED TO NETWORKED INFORMATION SERVICES

Reviewing recent literature about networked information services evaluation offered insight into methodologies, but it also indicated that such evaluation methodologies are less than fully developed. Research methodology literature in the areas of networked information services, government use of networked information resources, assessments of free-nets, and the six NSF/ARPA/NASA Digital Library research projects provided information of interest to the investigators.

McClure (1991) emphasized the need for user-based techniques rather than system-driven techniques for evaluating networked information services. These techniques take into account "the particular communication behavior, information use patterns, and work environments of potential users." McClure (1994) recommended four factors on

which to evaluate networked information services: extensiveness, efficiency, effectiveness, and impact. Specific techniques recommended were the use of focus groups, user logs, network-based data collection techniques, interviews, surveys, and site visits. Further research (McClure & Lopata, 1996), specifically in the academic networked environment, resulted in guidelines and suggestions that highlighted the value of using natural settings to more accurately assess the networked information service.

Networked information services, described by Bertot and McClure (1996), match the GILS environment in that there are multiple providers of the services, a range of information services available, growing use and access of the services, and a rapidly changing environment. Criteria for evaluating networked information services include service quality, usefulness, and the four factors previously cited by McClure (1994).

GILS is an example of a networked information service that occurs within a governmental setting. Bishop and Bishop (1995) highlighted the importance of user studies of networked information services for government accountability and effectiveness. They recognized that user studies need to reflect the complexity of human behavior and recommended new models for successful collaborations among users, social science researchers, and network decision makers.

User studies of free-nets also are of interest because these types of distributed networked information systems offer similarities to GILS. Newby & Bishop (1996) documented the methodology used to assess Prairienet in Champaign, Illinois. This report used descriptive statistics of web server transaction logs to identify characteristics of the users who access Prairienet. Patrick (1996) described the methodology used in a user survey of the National Capital FreeNet in Ottawa, Canada, which included a self-selected survey and a "random encouragement" survey.

Analysis of transaction log files offered another avenue for evaluation research. Noonan (1996) described the use of web usage statistics and listed four reasons for government agencies to be

interested in these sources of data. By analyzing web usage statistics, agency staff can demonstrate accountability, collect data to improve service, reach new audiences, and offer informative and useful means to disseminate information about the agency. The study offered the investigators practical guidelines for analyzing four common web transaction log files: access, error, referrer, and agent.

The six Digital Library research projects funded by the National Science Foundation (NSF), Advanced Research Projects Agency (ARPA), and the National Aeronautics and Space Administration (NASA) initiated a research stream helpful for evaluating distributed networked information services comparable to GILS. At the 1995 Allerton Institute, "How We Do User-Centered Design and Evaluation of Digital Libraries: A Methodological Forum," Bishop (1995) summarized the breadth of methodological issues addressed by the Digital Library Project researchers. She identified the data gathering techniques used by digital library (DL) researchers including log analysis, protocol analyses of user sessions, focus groups, in-depth interviews, user surveys, controlled observations with videotaping, collection of user comments and feedback, questionnaires, and written evaluations of testbed systems.

The Allerton Institute (1995) offered examples of research studies with methodological relevance to GILS evaluation efforts. At the Institute, Van House (1995) discussed user needs assessment and evaluation for the University of California - Berkeley's NSF/ARPA/NASA Digital Libraries Project. She identified three methodology areas which are "predecessor" in nature to digital library research: library evaluation with its focus on users' needs as the basis for evaluation, user-centered system design with its incorporation of user needs into system design, and usability analysis with its feedback methods.

Buttenfield's (1995) study, "User Evaluation for the Alexandria Digital Library Project," emphasized factors which researchers encounter when planning distributed network information services focusing on spatial data, which is a subset of material accessible through GILS. Methodological issues

for this project include targeting specific user classes, the lack of appropriate spatial metadata models, and a lack of understanding of user requirements. Both Van House and Buttenfield's work support methodological assumptions of the GILS evaluation project since the GILS evaluation also focused on user needs, on incorporating user needs into system design, on usability analysis, and on the need to target specific classes of users to determine user requirements.

The review of selected, recent methodology literature on evaluation of networked services clearly identifies such evaluations as an area under development. The investigators determined that the use of multiple methods to gather data is an emerging area of research methodology for evaluating networked information services. In addition, a focus on user needs is central in many of these studies. The research community is showing keen interest in developing new assessment strategies for evaluating networked information services.

3.2. AN EVALUATION FRAMEWORK FOR GILS

A multi-faceted information service such as GILS may be evaluated along different dimensions and from different perspectives. To accommodate the complexity of GILS, the investigators designed a framework that would guide a holistic approach to the evaluation. The framework identifies five interacting dimensions:

- **Policy:** policy goals and guidelines at both government-wide and agency levels that are shaping GILS
- **Users:** identification of user groups, their needs, their use of GILS, and their satisfaction with GILS
- **Technology:** technical implementation details including access mechanisms, implications of certain technology choices, and the effectiveness of that technology
- **Contents:** at the macro-level, what information resources are included in GILS; at the micro-level, the extent of

agency information resources described and the quality, degree of variance, accuracy, and usability of those descriptions

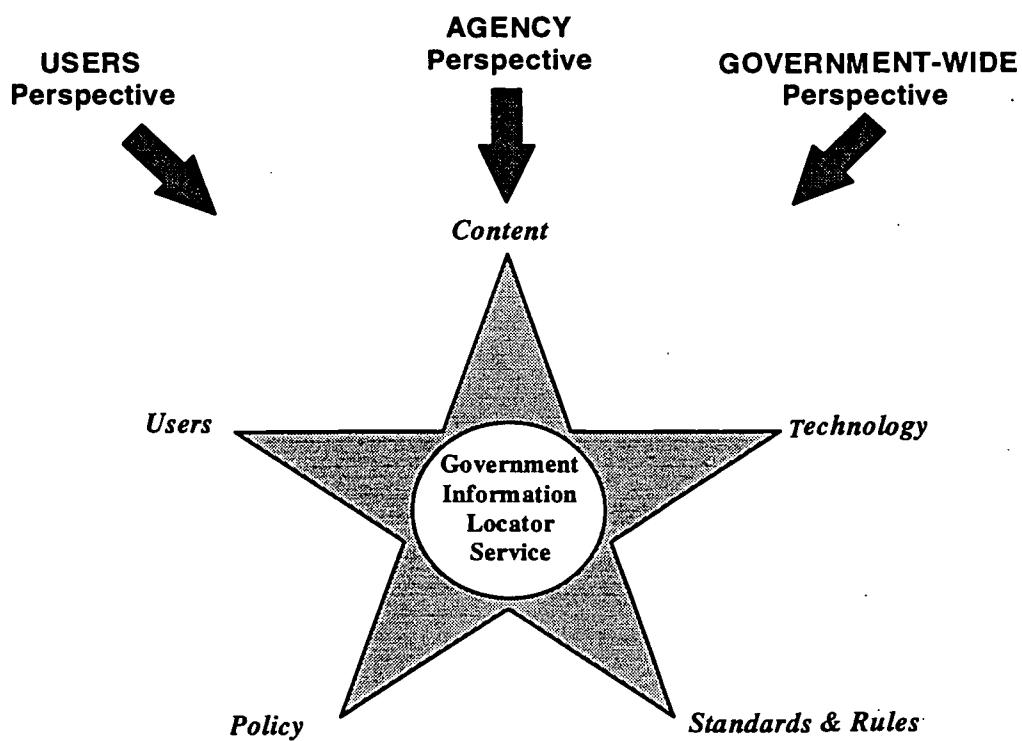
- **Standards and Rules:** utility of standards to ensure consistency in GILS information, to offer broader connection, access, and retrieval of information.

The evaluation framework also includes three perspectives, representing the "views" of various stakeholders in GILS: Users, Agency, and Government-Wide. The three perspectives helped

to focus the evaluation on the need to represent different views held by different stakeholders during implementation and use of a networked-based information service. The investigators were also aware that the study findings would be of interest to people viewing GILS from these various perspectives.

Together the three perspectives and the five dimensions capture the complexity of GILS as a networked information service and guided the research design and data collection activities. Figure 3-1 presents the evaluation framework.

Figure 3-1
Framework for GILS Evaluation:
Perspectives and Dimensions



3.3. EVALUATION GOALS AND STUDY QUESTIONS

The *Technical Proposal* (Moen & McClure, 1996a) enumerated the following goals for the study:

- Examine and describe how GILS is serving users in locating and accessing government information
- Examine and describe agencies' GILS implementation experiences
- Identify and document success factors and/or barriers affecting agencies' GILS implementations
- Examine and describe agencies' use of GILS as an information resources management tool
- Determine if changes to the GILS policies or technical specifications are needed to make it a more useful tool for agency information resources management
- Provide recommendations and strategies that will assist agencies to improve their GILS applications.

At the outset of the evaluation study, the investigators identified a number of study questions, derived from the project goals, to guide initial information gathering and data analysis activities of the project:

- Who are current GILS users?
- How is GILS serving users to locate and access government information?
- What have been agencies' experiences in implementing GILS?
- What are the critical success factors affecting agencies' GILS activities?
- What are the barriers affecting agencies' GILS activities?
- What are GILS "best practices" that could be useful for all agencies?
- How are agencies using GILS as an information management tool?
- To what extent are agencies conforming to *FIPS Pub. 192* (for structure and contents of locator records, and for making their records available via Z39.50)?

- What changes are needed either to GILS policies or technical specifications to improve the utility of GILS for users, agencies, and the Federal government?

The research strategy assumed that the study questions might be refined and modified as the investigators collected and analyzed data. Addressing these and similar study questions, however, helped the researchers link activities directly to project goals.

In the initial stages of research design, the investigators identified research activities that could be used to collect and analyze data needed for an understanding of the current status of GILS. Not all possible activities and techniques originally considered during the design of the study became part of the research activities (e.g., in the technology and standards dimensions of the framework, one activity considered was to test Z39.50 protocol level compliance of the GILS implementations). Resource constraints forced the investigators to select only significant and cost-effective data collection activities that would best serve the purposes and goals of the study.

3.4. THE RESEARCH STRATEGY

The investigators positioned this study within a qualitative, naturalistic research context (Maxwell, 1996; Patton, 1990), although the study also used quantitative techniques and mixed qualitative and quantitative methods (Creswell, 1994). The qualitative context recognizes the evaluation's emphasis on process over measurement, understanding and learning over hypotheses testing. A qualitative research approach was appropriate to produce a richly detailed, holistic understanding of GILS. Moreover, it also allowed study participants to better describe their experiences and use of GILS.

The research strategy provided overall direction to accomplish the purpose and goals of the study. This strategy involved the use of a variety of research techniques and methods, including:

- Site visits
- Focus groups
- Survey
- GILS record content analysis
- Scripted online user assessments
- Web server transaction log analysis
- Policy and literature review.

As the research progressed, the investigators modified specific research techniques and methods to reflect their understanding of GILS.

Each technique and method required the development and testing of procedures and instruments. In some cases, procedures could be lodged within existing and well-known data collection activities. For example, the investigators pre-tested questions and the survey instrument for focus groups and the survey respectively. For three of the techniques, the investigators developed new, exploratory procedures, instruments, and analysis procedures that had not been well-articulated in either the literature or previous research by others. For example, in the case of the record content analysis, no pre-existing operationalization of "GILS record quality" existed.

Early data collection techniques provided the investigators with data that was analyzed and informed subsequent data collection including the HTTP transaction log analysis, scripted online user assessment, and GILS record content analysis. The investigators then analyzed and synthesized data from these data collection techniques to create preliminary findings. Member checks, follow-up interviews, and discussions among the investigators enriched the preliminary findings and served to further the trustworthiness of the data and the findings. As part of the investigators' concern for trustworthiness, they pre-tested data collection instruments including focus group questions, site visit protocols, and conference survey questions. Figure 3-2 presents an overview of the GILS research design.

The research design was structured yet flexible in allowing refinement of questions and modification

of data collection techniques based upon initial data collection and analysis. Early data collection and analysis informed choices regarding subsequent data collection and analysis. The research design was necessarily evolutionary.

3.5. MULTIPLE METHODS AND SPECIFIC DATA COLLECTION TECHNIQUES

Since an evaluation of GILS needed to examine diverse factors (e.g., nature and type of resources to be described by locator records, agency resources available, etc.), the investigators needed diverse but complementary data gathering techniques to capture as fully as possible the breadth and depth of issues. The investigators matched research information needs (e.g., information needed about each dimension of GILS, and information needed to answer study questions) with appropriate quantitative and qualitative research techniques (Creswell, 1994).

The investigators selected and utilized one or more methods on the basis of satisfying the information needs of each component of the study. As an example, site visits to agencies allowed the investigators to interview agency staff to fully realize all aspects of an agency's usage and implementation experiences with GILS from various participant perspectives. In a parallel manner, focus group sessions with various types of GILS stakeholders represented opportunities for the investigators to bring together homogeneous groups of stakeholders to represent common-interest perspectives.

These methodologies used theoretical rather than statistical sampling. Unlike the latter, which is designed to provide data subject to statistical verification, theoretical sampling allowed capture of incidents of difference, and, in a progressive fashion, built a broad foundation for subsequent analysis and understanding (Glaser & Strauss, 1967).

Figure 3-2
Overview of GILS Evaluation Research Methodology

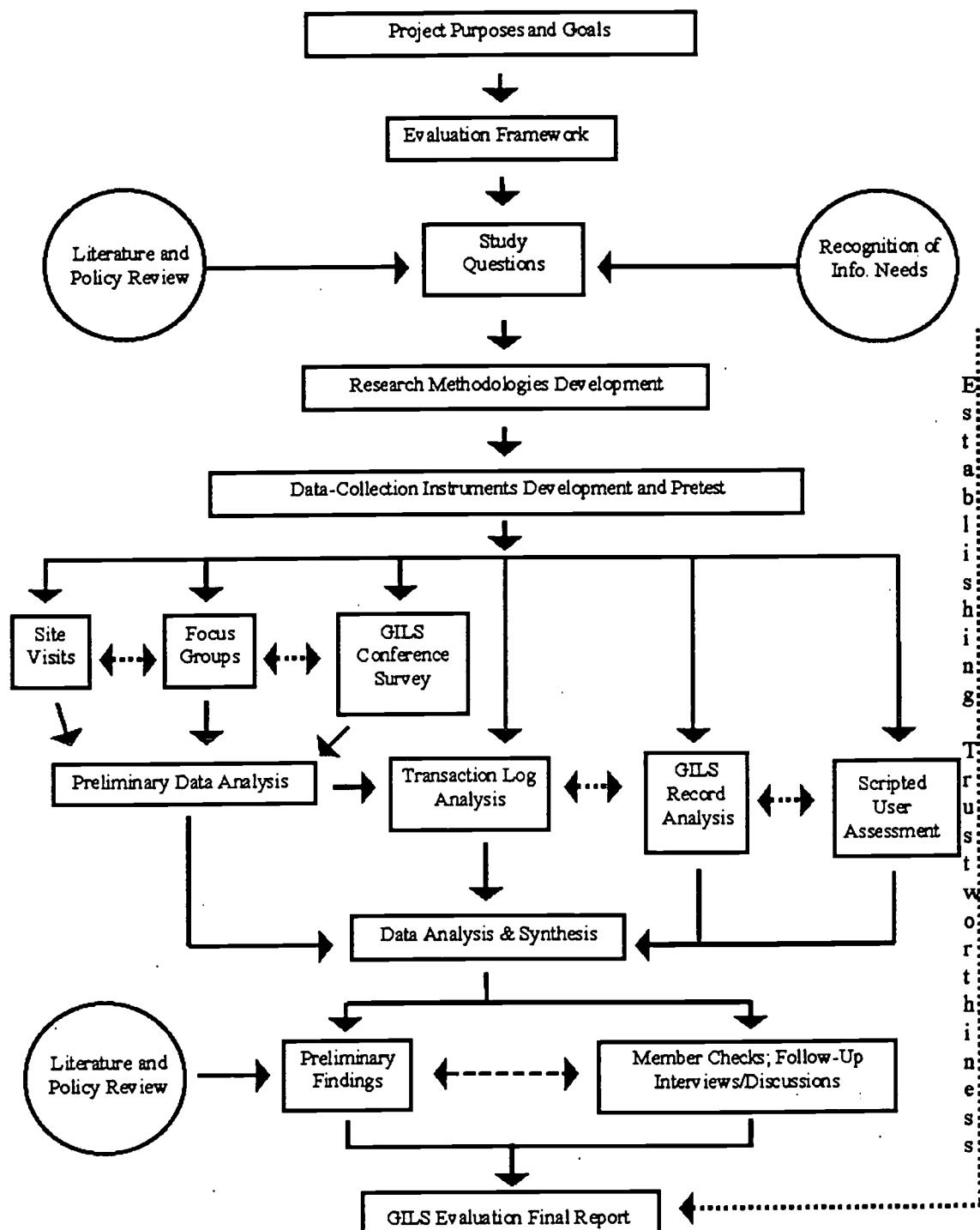


Table 3-1
Data Collection Techniques

Technique	Primary Method	Information Obtained	Forms of Data	Analysis
Site Visits	Qualitative	Agency-specific experiences as described by various agency staff as participants in GILS	Narrative Text	Content Analysis
Focus Groups	Qualitative	Stakeholder-specific perspective on GILS	Narrative Text	Content Analysis
Survey	Quantitative	Quantifiable assessments of key GILS issues	Numeric Data	Descriptive Statistics
GILS Record Content Analysis	Quantitative	Measurement and assessment of GILS record quality	Numeric Data	Descriptive Statistics
Scripted Online User Assessments	Qualitative Quantitative	User's assessments of GILS as a networked service	Narrative Text Numeric Data	Content Analysis Descriptive Statistics
Web Server Transaction Log Analysis	Quantitative	Machine-generated data of users' interaction with GILS	Numeric Data	Descriptive Statistics
Policy and Literature Review	Qualitative	Analysis of the policy environment and specific policies providing the context for GILS	Narrative Text	Content Analysis

Table 3-1 summarizes the data collection techniques used in the study. Each technique is associated with one or more primary methods (i.e., qualitative or quantitative), the kind of information obtained, and the form of the resulting data. Study team members then analyzed the resulting data using appropriate analytical techniques.

The following briefly describes each technique and how it was used in the evaluation. Each description includes a summary of number of people involved, activities carried out, etc. For additional detail on each technique, see Appendices C-1 through C-6.

3.5.1. Site Visits

Investigators conducted one-day visits to agencies to observe specific environments of GILS implementation (see Appendix C-1). The following is a list of agencies selected for site visits and the dates of occurrence:

- Department of Defense, Defense Technical Information Center (November 15, 1996)

- Environmental Protection Agency (October 23, 1996)
- Government Printing Office (November 15, 1996)
- Department of Treasury (January 10, 1997).

Investigators carried out guided interviews with personnel from many administrative and functional areas. Site visits also included one focus group of agency staff, examination of relevant agency documentation, and tours/demonstrations. Site visits provided detailed understanding from participants' perspectives of agency GILS implementation issues. A total of 46 agency staff participated in the site visits.

Through interviews with knowledgeable agency staff (i.e., policymakers, managers, systems/technology staff, intermediaries, librarians, records managers, and agency end users), the investigators collected data to understand and describe agencies' GILS efforts. The agency site visits enabled the investigators to examine policy, management, technology, and human aspects of

agencies' implementation experiences. Finally, these site visits provided data for detailed case study description of experiences.

The investigators also conducted two additional sets of interviews and discussions during the study. Investigators met with two individuals involved with the planning of the Advanced Search Facility to learn about that technology initiative. Staff at the National Archives and Records Administration (NARA) met with the investigators to discuss GILS-related activities undertaken by NARA (i.e., training and development of record creation guidance). The discussion at NARA also focused on GILS and records management issues that had surfaced during the study.

3.5.2. Focus Groups

Investigators conducted a series of "carefully planned discussion[s] designed to obtain perceptions on a defined area of interest in a permissive, non-threatening environment" (Krueger, 1988). Focus groups brought together groups of stakeholders, allowing individuals with common interests an opportunity to explore shared beliefs and goals with respect to GILS (see Appendix C-2). A total of 83 people participated in these focus groups. The following lists the focus groups and dates of occurrence:

- North Texas area Government Documents Librarians (October 31, 1996)
- Public Interest/Public Access Stakeholders (November 13, 1996)
- State/Local GILS Implementors Stakeholders (November 13, 1996)
- Vendor/Technology Stakeholders (November 13, 1996)
- Future Issues Stakeholders (November 14, 1996)
- Records Managers Stakeholders (November 14, 1996)
- Small Agency Council (February 13, 1997).

These sessions provided information about the knowledge and awareness of GILS by important

stakeholder groups and an opportunity to document their expectations and their encounters with GILS. The intent of this data collection activity was to understand users' impressions, understanding, expectations, satisfaction, and frustrations with the current implementations of GILS. As part of the analysis of this data, the investigators identified user requirements that are and are not being met. Further, the investigators were able to compare original objectives of GILS with what users expect today from a government information locator.

3.5.3. Survey

Investigators developed a survey instrument administered to participants of the GILS Conference in November 1996. Respondents assessed key GILS policy and other issues on the questionnaire. The survey also provided assessments of conference participants' knowledge of GILS policies, attitudes, and experiences as well as qualitative information concerning expectations and lessons learned. A total of 181 conference participants completed the survey (see Appendix D-3 for a copy of the instrument and Appendix C-3 for methodology).

3.5.4. Record Content Analysis

Investigators developed a procedure for analyzing the content of GILS records through an examination and assessment of a sample of GILS records from known GILS agency implementations (see Appendix D-4 for a copy of the instrument and Appendix C-4 for methodology). The investigators employed specific tests to operationalize a set of criteria that included accuracy, serviceability, completeness, and currency.

The intent of this research activity was twofold:

- Develop criteria and procedures for assessing GILS records for use in the study and for subsequent adaptation by agencies for their ongoing assessment of record quality.
- Isolate possible trends in GILS record character and quality of the population of GILS records (approximately 5,000).

Investigators selected a sample of 80+ GILS records from 42 agency GILS implementations and examined and coded approximately 4500 data instances in the records for assessment.

3.5.5. Scripted Online User Assessment

Investigators developed an exploratory method of scripted online user assessments to generate an understanding of user expectations for and reactions about GILS (see Appendix D-5 for a copy of the script and Appendix C-5 for methodology). In this data gathering activity, the investigators sought to understand GILS from the perspective of users. The investigators had developed scripted online assessments of networked information resources (e.g., government bulletin boards, and network services) for previous studies (see Bertot & McClure, 1994; Bertot & McClure, 1996a, 1996b). This methodology is exploratory, and its use in the GILS evaluation provided opportunity for enhancement and refinement.

The investigators developed a tightly scripted set of browsing, searching, and retrieval tasks that highlighted salient features targeted for encounter by users. The goal of this scripted assessment was to draw from users the extent to which they understand what GILS is, whether their expectations of GILS are in line with how GILS has been implemented, and to lead to a set of user-based requirements for improvements to GILS. Ten undergraduate and graduate students participated in the assessment.

Like other methods used in the evaluation study, the scripted online user assessment served several purposes. One was to develop and test reliable scripts and assessment procedures suitable for use by agencies in evaluating their specific implementations. The method also informed the investigators' understanding of GILS from the perspective of users.

3.5.6. Web Server Transaction Log Analysis

GILS implementations currently use a base of technology that includes Web browsers, HTTP and

Z39.50 servers. A benefit of GILS implementations using the Web is the availability of a set of common HTTP server transaction log files (Noonan, 1996). The investigators incorporated an exploratory transaction log analysis to assist in the assessment of GILS. In the past, transaction log analysis research has produced methods and procedures for understanding user interaction with online systems. Log analysis in the networked environment, however, is an emerging area of research.

Investigators developed a set of procedures and analyses to examine data from one agency's web server transaction log files. The procedures generated data for statistical analysis of user transaction activity on an agency's GILS server (see Appendices C-6 and E-4).

Using sample data from a single agency's HTTP log files, the investigators tested and refined procedures for statistical analysis of user transactions. Additionally, the investigators developed procedures to discern patterns in user interaction with the Web and GILS information spaces.

The primary intent of this activity was to develop and test new tools for log analysis. The result of this effort is a set of analysis procedures that agencies can use in ongoing assessment of their GILS implementations.

3.5.7. Policy and Literature Review

Investigators completed a review of GILS policy instruments, regulations, laws, and related literature to provide an understanding of the current environment that is the context for GILS implementations (see Chapter 2). This review enabled the investigators to develop recommendations for changes and enhancements to policies—both government-wide and for individual agencies.

This research activity identified the current policy environment for GILS as a basis for synthesizing policy prescriptions and describing GILS in the evolving policy environment. Such analysis was central to clarify and understand the policy context that affects design, management, implementation,

and use of GILS. In addition, the researchers collected and examined agency documents for examples of beneficial and transferable policies. Recommendations could then be offered to clarify, expand, or revise the policy framework for improved coherence and understanding.

3.6. MULTIPLE DIMENSIONS/ MULTIPLE DATA EVENTS

One or more of the techniques described above collected data related to each of the dimensions in the evaluation framework (Figure 3-1). The following sections describe briefly the scope of each dimension and identify data collection activities associated with its study.

3.6.1. Technology

The dimension of technology included technical implementation details such as access mechanisms and implications of certain technology choices by Federal agencies and policymakers. Data collection to explore the technology dimension featured:

- Focus group of information technologists and vendors
- Site visit interviews with information resources management (IRM) staff at several Federal agencies
- Transaction log analysis of agency data
- Interview with Advanced Search Facility (ASF) staff
- Content analysis of GILS records.

At the 1996 GILS conference, the investigators invited vendors and technologists to a focus group session to discuss both existing and future technology options for GILS. This context brought together a group of stakeholders whose views on GILS technology included market potential, feasibility and desirability of future technological developments, and an evaluation of GILS functionality from a group of technology-informed users.

Site visits with IRM staff at selected Federal agencies enriched data gathering through use of personal interviews. Within different agencies, IRM and systems staff fulfilled a variety of roles as part of the process of implementing GILS as a networked information resource. Investigators interviewed those agency staff who guided the GILS technical efforts. These interviews aided in an understanding of key issues, challenges, and critical success factors for the agency.

An additional data gathering technique included an exploratory log analysis activity designed to assist in the evaluation of GILS usage. Transaction analysis of log files from an agency's GILS records provided the investigators with an important tool for understanding usage of a networked-based information service.

3.6.2. Content

The dimension of content, at the macro-level, identified the information resources included or covered in GILS, and at the micro-level concerned the quality, degree of variance, accuracy, and usability of the information resource descriptions covered by GILS. Data gathering techniques for this dimension included:

- Content analysis of GILS records
- Site visit interviews with record creators and others
- Focus group sessions with users, records managers, librarians, and others
- Survey questions related to scope and coverage of GILS records
- Scripted online user assessment.

Investigators developed criteria and assessment methods to evaluate a sample of GILS records. National Archives and Records Administration (NARA) *The Government Information Locator Service: Guidelines for the Preparation of GILS Core Entries* (National Archives and Records Administration, 1995a) provided a basis for the development of the criteria. Agency GILS implementors used these guidelines in creating agency GILS records. To understand implementors'

decision-making with regard to record content, site visits to agencies included interview sessions with record creators. These interviews with the staff who had personal involvement in the record creation process contributed important information on the strategies which shaped decisions about an agency's GILS records. Focus group sessions, survey questions, and user assessment also provided the investigators with perceptions and perspectives on the usefulness and value of GILS records from different user groups.

3.6.3. Users

The user dimension concerned identification of GILS users: their needs, their usage of GILS, and their satisfaction with GILS. Data gathering techniques for this dimension included:

- Scripted online user assessment
- Focus groups with government document librarians, records managers, and public interest groups
- Site visit interviews with agency staff who are GILS users, and with intermediaries
- Focus group with state and local GILS implementors
- Focus group with records managers
- Survey.

GILS users are not a homogeneous group, but rather consist of a variety of separate user groups including librarians, public citizens, records managers and other staff members at the implementing agencies, and state and local GILS implementors.

Agency site visit interviews included discussions with staff to learn about that agency's efforts to involve users in the agency's planning activity and the agency's experiences with public use of GILS as an effective means to obtain government information. Site visit interviews with agency staff who directly supported public access to government information also provided information on users' perceptions of GILS. A number of the focus groups gathered information about specific groups of users such as records managers, librarians, and public

interest groups. The scripted online user assessment collected data on users' interaction and response to specific GILS implementations.

3.6.4. Policy

The policy dimension of the evaluation framework described the policy environment for U.S. Federal GILS implementation. Data gathering events and activities for this dimension included:

- Policy review and analysis
- Focus group with Federal information policymakers (included in Future Issues Stakeholder session)
- Site visit interviews with agency policymakers
- Survey.

Investigators conducted a policy review of legislation, executive orders, and other guidelines which represented formal information policy with respect to GILS. The review highlighted key policy issues as well as identified changes in policy since GILS' inception in 1994.

Focus group sessions with Federal information policy stakeholders and site visit interviews with agency policymakers provided opportunities for important stakeholder groups to not only inform the investigators as to current and future policy goals in this area but also to share among themselves mutual insights and concerns. Site visit interviews enabled the investigators to gain an understanding of agencies' internal policy with respect to networked information resources. Investigators obtained and analyzed agency policies and guidelines when available. Finally, the survey included questions about respondents' familiarity and understanding of information policy sources for GILS as well as assessment of existing policy guidance.

3.6.5. Standards and Rules

The standards and rules dimension addressed the utility of standards to ensure consistency in GILS information, and the use of this means to support broader connection, access, and retrieval of

information. Data gathering techniques for this dimension included:

- Site visit interviews with agency staff about the role of standards and the use of Z39.50 in GILS
- Site visit interviews with agency staff regarding the general awareness and use of NARA guidelines for record creation
- Content analysis of GILS records
- Survey
- Focus group sessions with users about the value of standards and the general level of awareness of standards with respect to GILS.

The investigators interviewed administrators and IRM staff at Federal agencies to learn of their general awareness of standards and specific use of Z39.50 within that agency's implementation. The survey included questions designed to elicit respondents' awareness and usage of standards.

It is important to note that the five dimensions of the evaluation framework and the multiple data collection techniques did not exist in isolation from each other. Multiple data collection techniques not only enabled the investigators to explore aspects of any one dimension from a variety of perspectives but also provided for exploration of the relationships and interaction of these dimensions. The combination of study activities resulted in an integrated and carefully constructed view of U.S. Federal GILS implementation.

3.7. AN INTEGRATED VIEW OF GILS

The research strategy developed for the assessment and evaluation of GILS incorporated multiple methods and techniques to arrive at a holistic view of GILS and to address the study questions posed at the outset. These data collection and analysis activities provided macro-, midlevel-, and micro-views of GILS. The macro-view allowed the investigators to examine broader policy and organizational issues related to GILS. At the midlevel, user groups provided insight into their understanding and expectation of GILS. And at the

micro-level, the record content analysis and individual scripted online user assessments identified and assessed discrete aspects of GILS that informed recommendations on improving GILS. When combined and synthesized, these views allowed the investigators to gain a holistic understanding of many aspects of GILS.

3.8. CONCLUSION

For this evaluation study, the investigators devised an innovative research approach to explore the multi-faceted nature that we assert is not only characteristic of GILS but of other complex networked information services. The investigators also developed and enhanced specific user-based data collection techniques for the evaluation and combined these techniques in effective ways to understand and evaluate the current state of GILS implementations. User-based assessments can be a countervailing force to the glamour and hype of the

sophisticated technology that provides such vital ways of organizing and accessing information in the digital age.

Both the number and array of data-gathering techniques employed by the investigators produced not only an integrated set of wide-angle and zoom "snap shots" of GILS but also a set of procedures that can be useful to agencies when assessing their own GILS implementations (see Appendices C-1 through C-6). The investigators hope that information policymakers as well as networked services implementors will build on and refine the procedures specified for record content analysis, transaction log analysis, and scripted online user assessments to serve tactical and strategic objectives for information resource management.

The evaluation literature addressing digital libraries reflects the need for multi-method and multi-level assessment of complex networked information services (Bishop, 1995). GILS is also a complex information service existing within the larger networked information infrastructure. The findings reported in Chapter 4 underscore the complexity of the implementation, coordination, and utility of networked information services.

Chapter 4

Findings and Recommendations

4.0. INTRODUCTION

Since the publication of the Office of Management and Budget (OMB) Bulletin 95-01 in December 1994, Federal agencies have undertaken implementations of the GILS Profile to comply with the mandate of the Bulletin. The study identified approximately 45 agencies that have some form of GILS implementation. Study results indicate that for a handful number of agencies, GILS has improved access to and knowledge of agencies' indexes, catalogs, finding tools, and other "metadata sources."

Yet, for many other agencies, the Federal GILS initiative has been little more than another unfunded mandate that received little administrative support, has not met original objectives, has provided few benefits to agencies and users, and has little visibility either in government or with the public. Further, much confusion exists over what GILS is and should be. The agency GILS that are operational have limited use and study participants assessed them as difficult to use. In addition, the records management component of GILS mandated by OMB Bulletin 95-01 was poorly conceived, and GILS as a records management tool does not assist records managers in meeting their responsibilities related to records management.

For the majority of agencies, their GILS effort and expense has not resulted in adequate or tangible benefits—regardless of how one defines "benefits." Many agencies reported that limited resources were available for GILS, and the lack of resources and effort by some agencies (e.g., one agency had created only a single GILS record) has limited the potential utility of GILS as a government-wide information locator. Nonetheless, the study also finds that agencies and users are positively disposed to the concept of GILS, defined by OMB Bulletin 95-01 as a service that "will identify information resources throughout the Executive Branch, describe the information available, and provide assistance in how to obtain the information" (Office of Management and Budget, 1994b). With a conscientious refocusing, GILS could have great potential to improve access to and use of Federal government information.

The investigators conclude that GILS—as a concept and mechanism—has an important role to play in discovering and accessing government information in the networked environment. The investigators affirm the underlying architecture of GILS: standardized metadata records, decentralized agency-based locators, standard protocols (i.e., Z39.50) for intersystem information retrieval. The U.S. GILS implementation, however, has not achieved fully the vision of a “virtual card catalogue” of government information nor have the agency GILS implementations matured to the extent of providing a coherent and usable government-wide locator service. The investigators conclude that many of the current shortcomings with GILS relate to problems of focus, scope, and administration rather than a fundamental flaw in the concept of GILS.

The investigators recommend the Federal GILS initiative be refocused to clarify both purpose and functions of GILS implementations. A refocused GILS initiative can assist in providing guidance to all agencies as they continue their implementations as well as offering clearer evidence of the utility of GILS to the many agencies that have concluded GILS is neither useful nor beneficial.

A refocusing of the GILS effort provides the next evolutionary step for U.S. GILS development. It will build upon the work accomplished and upon the experiences and lessons learned for improving public access to government information in the networked environment. Policymakers, however, must draw a clear line of demarcation between the early GILS implementation period (i.e., 1995-1996) and a refocused and reengineered GILS. This line of demarcation is essential because it represents an acknowledgement by policymakers and implementors that:

- Many agencies are now unwilling to put additional resources into an initiative of questionable utility
- Lessons have been learned by policymakers and implementors from the early implementation experience
- The refocused GILS will address shortcomings and issues made visible from existing implementations.

One way in which a refocusing of GILS can be underscored is through a change in the name to reflect, for example, a “second release” of the U.S. Federal GILS service.

These general statements of findings and recommendations are detailed in the subsequent sections of this chapter. The chapter has two opening sections that describe the Federal context in which the GILS initiative occurred and the current status of agency GILS implementations. The chapter then organizes findings and recommendations into four primary opportunities which are discussed in Sections 4.3. through 4.7.:

- Refocus GILS for Clarity of Purpose and Utility
- Improve GILS Efficacy in Networked Information Discovery and Retrieval (NIDR)
- Resolve GILS Relationships with Other Information Handling Functions and Processes
- Increase GILS Awareness.

The opportunities provide policymakers and implementors with a framework for addressing areas where the Federal GILS initiative can be improved. Each section in this chapter that describes one of these opportunity sections includes a table identifying relevant findings, recommendations, and supporting sources of evidence. Table 4-1 summarizes the four opportunities and associated findings and recommendations.

The recommendations reflect the investigators' analysis, synthesis, and understanding of the data collected during the study and the findings reported here. A number of the findings and issues uncovered during the study, however, presented challenges in devising specific recommendations. This is especially the case in recommendations relating to the second opportunity area, “Improve GILS Efficacy in Networked Information Discovery and Retrieval (NIDR).” NIDR is an active research area; researchers and early implementors have recognized the complexity of many NIDR problems in the past several years. For example, in the areas of metadata and distributed search and retrieval,

Table 4-1: Opportunities, Findings, and Recommendations

Opportunity: Refocus GILS for Clarity of Purpose and Utility	
Findings	
4.3.1. People Are Confused about GILS Mission, Purposes, and Uses	
4.3.2. Expectations for GILS Are Evolving	
4.3.3. Government-Wide Administrative Coordination and Policy Oversight Are Lacking	
4.3.4. Smaller Agencies Feel Special Burden and Frustration	
4.3.5. Agencies' Cultures and Missions Promote Different Commitment to GILS	
4.3.6. Intra-Agency Efforts Reflect Different Levels of Enthusiasm for GILS	
4.3.7. GILS Benefits Compared to Burdens Are Not Clear	
Recommendations	
4.3.8. Focus on Public Access to Government Information	
4.3.9. Focus Scope of Descriptions On Network-Accessible Information Resources	
4.3.10. Identify Responsibilities and Authority for Policy Leadership, Government-Wide Coordination, and Oversight	
4.3.11. Implement a Refocused GILS Initiative	
4.3.12. Require Agency Reporting on GILS Progress and Reward Agencies That Achieve Stated Objectives	
4.3.13. Ensure Ongoing, User-Based Evaluation for Continuous Improvement	
Opportunity: Improve GILS Efficacy in Networked Information Discovery and Retrieval (NIDR)	
Findings	
4.4.1. Web Technology Has Raised Questions about the Role of GILS	
4.4.2. GILS is an Agency-Centric, Rather than Government-Wide, Service	
4.4.3. GILS Metadata Are Difficult to Capture	
4.4.4. Limited Updating and Maintenance of GILS Records	
4.4.5. No Clear Agreement on Adequacy of GILS Record Data Elements	
4.4.6. Different Types of Resources Represented in GILS Records	
4.4.7. User Reaction to GILS Is Not Positive	
4.4.8. GILS Record Display Varies Widely and Is Criticized by Users	
4.4.9. User Orientation and Instruction is Inadequate	
Recommendations	
4.4.10. Continuously Evaluate GILS Policies and Standards against Emerging Technologies, Especially the Web	
4.4.11. Specify Resource Types And Aggregation Levels	
4.4.12. Enforce Consistent Use Of Metadata That Are Empirically Demonstrated to Enhance NIDR	
4.4.13. Improve Presentation of Metadata	
4.4.14. Develop Policy and Procedures for Record Maintenance	
4.4.15. Promote Interagency Cooperation and Use of GILS for One-Stop Shopping Functionality	
Opportunity: Resolve GILS Relationships with Other Information Handling Functions	
Findings	
4.5.1. GILS Does Not Support Records Management Activities	
4.5.2. GILS Relationship with Agencies' Inventories of Information Resources Is Not Clear	
4.5.3. GILS Relationship with FOIA and EFOIA Is Unclear	
Recommendations	
4.5.4. Uncouple the Refocused GILS—as an Information Discovery and Access Service—from Records Management	
4.5.5. Derive GILS Metadata from Other Information Handling Processes	
Opportunity: Increase GILS Awareness	
Findings	
4.6.1. No Program for GILS Promotion and Education Exists	
4.6.2. Potential User Communities Lack Familiarity with GILS	
4.6.3. GILS Usage Is Limited	
Recommendations	
4.6.4. Develop and Formalize GILS Promotion, Education, and Training Strategies	

there are prototype implementations but complete and scalable solutions will await additional research (see Lynch, 1997; Lynch, et al., 1995).

The investigators believe that the recommendations will contribute to determining the next evolutionary steps for the U.S. Federal GILS initiative. Chapter 5 proposes a framework of action that identifies next steps for a refocused GILS effort. Ultimately, however, it is the project's advisory group, the sponsoring agencies, and the GILS Board that must determine what is to be done with GILS. The findings and recommendations reported here can provide substance as well as points of departure in the deliberations of the advisory group, Federal policymakers, implementors, and the GILS Board.

In keeping with the charge of the study to examine how the GILS initiative serves users (see Moen & McClure, 1996a, for the study's *Technical Proposal*), the findings reported here rely on data collected from the various groups of "users" involved with GILS. The term "users" of GILS belies the complexity of identifying who, specifically, the GILS users are. For purposes of this discussion, user groups appear to be best described in the following terms:

- **Federal agency staff:** including agency GILS implementors, agency managers, records managers, policymakers, agency librarians, and others
- **State and local government staff:** including state and local GILS implementors, state library agencies, records managers, librarians, and others
- **Non-governmental individuals:** including librarians, public advocacy groups, journalists, the "public," those with special subject interests, and others.

The user-based evaluation designed by the investigators recognized and valued the various special interests and perspectives of all these user communities.

Chapter 3 discussed the multi-method approach used in this study (with complete details of the specific methods in Appendices C-1 through C-6). The data collection and analysis activities carried out during

the study produced a significant amount of information from which the study's findings and recommendations flow (Appendices E-1 through E-4 contain detailed results organized by data collection activity). The findings reported here are based on data collected through the following sources of evidence:

- Site visits
- Focus groups
- Survey
- GILS record content analysis
- Scripted online user assessments
- Web server transaction log analysis
- Policy and literature review.

For each of these activities, the study team compiled results and produced detailed summaries. For example, the summary for a typical site visit is about 25–40 pages plus appendices. Particular findings may be based on data produced from one or more of the study activities. Often, similar findings emerged from more than one data collection effort. Instead of reporting the results for each data collection activity or instrument, this chapter organizes the findings and recommendations into opportunity areas. When appropriate, the discussion links evidence from specific sources or assessment activities to particular findings.

4.1. GILS IN THE FEDERAL CONTEXT

OMB Bulletin No. 95-01, issued in December 1994, formalized the U.S. Federal GILS initiative and provided policy guidance for its implementation. At the same time, the National Institute for Standards and Technology (NIST) released *Federal Information Processing Standard (FIPS Pub.) No. 192* that provided technical specifications and implementation guidelines in the GILS Profile. In February 1995, NARA published *The Government Information Locator Service: Guidelines for the Preparation of GILS Core Entries* to assist Federal agencies in the creating GILS records.

Approximately 2 years have passed between the formal announcement of the initiative and this

assessment effort. As a government-wide initiative, GILS is relatively young.

During this 2-year period, a number of key factors affected the Federal government environment and the GILS initiative. First, the government launched GILS during a time of significant downsizing, budget cutting, and reorganization of the Federal government. There was substantial discord between Congress and the Administration regarding the appropriate role and size of many government agencies. Agencies, oftentimes, were under pressure to reduce budgets and reduce staff size, yet also expected to demonstrate greater productivity and "streamline" operations. The closing of the Federal government due to budget disagreements between Congress and the Administration in 1996 also contributed to an already difficult work environment (in fact, the shutdown caused an extension to the first OMB Bulletin 95-01 deadline for GILS implementations). In short, agency morale often suffered.

Second, recent years have seen, perhaps, the greatest amount of activity related broadly to information management issues, policies, and legislation in the history of the U.S. Federal government. *The Government Performance and Results Act of 1993, The Paperwork Reduction Act of 1995, The Information Technology Management Reform Act of 1996, and The Electronic Freedom of Information Act Amendments of 1996*—to name but a few legislative initiatives—significantly affected the environment of information management in the Federal government (see the policy review section in Chapter 2).

In addition to legislation, a host of policy issues related to encryption, privacy, information technology (IT) procurement, standards, electronic records management, access to government information, the National Information Infrastructure (NII), revision of Federal printing laws (e.g., Title 44 USC), and other topics required the attention of agency information managers. Between legislated mandates and other information management/policy issues, there has been no lack of work or policy issues demanding attention from agency officials in the broad area of information management.

Finally, the emergence of new IT and related applications has also been significant. Since early 1994, Internet—specifically Web—applications have dominated and redefined access to and dissemination of information. Due in part to initiatives related to the NII and the Administration's interest in utilizing Internet and Web technology, agency use of Web applications for disseminating information and providing electronic information services grew exponentially. One need only examine GPO Access, NTIS' FedWorld, the Library of Congress' Thomas legislative search system, and the many agencies that have established Web sites in the past several years to gain an appreciation for the use and interest in Web technology by the Federal government. The well-known Federal Web Locator maintained at The Villanova Center for Information Law and Policy <<http://www.law.vill.edu/Fed-Agency/fedwebloc.html>> now indexes and provides access to nearly 1,000 Federal Web sites. Truly, the Web changed fundamentally the ways many agencies use the Internet for presenting and publishing information.

The development and use of the Internet and Web technology by Federal agencies is a significant factor that shaped the Federal information management environment in recent years. At the time of the work on the technical and policy specifications that underlie the GILS initiative (1993–1994), the growth and development of Web-based services could not have been foreseen. The Web phenomenon was a surprise to almost everyone, including the designers and developers of the GILS Profile as well as U.S. Federal GILS policymakers. To some degree, the GILS initiative may have been swallowed by Web developments. The latter clearly caught the interest of both the public and government officials much more so than GILS because the Web was concrete and real—people could see it, use it, and understand its potential. The Web now offers agencies a mechanism for easy electronic publishing and dissemination of large amounts of information, and users can access the full-text of documents.

GILS as a set of metadata records describing government information—or GILS as an implementation of Z39.50—is not nearly as glamorous nor easily understood as the Web. While the Web offers new opportunities to agencies, it has a

limited capability to help users discover and locate government information resources, especially on a government-wide basis. GILS metadata records and the use of Z39.50 as a standard mechanism for interoperable search and retrieval across GILS databases, however, has the potential for solving the problems of information discovery in the networked environment.

Several other factors will be identified in this chapter that affected the success of GILS as implemented by Federal agencies. But these three—downsizing government, expanding information management legislation and policy issues, and Internet/Web development—should be recognized as affecting the current status of the U.S. Federal GILS initiative. As a long-time information resources management (IRM) official commented to one of the investigators during the study, “never in my years working for the government have I seen as much change in information management and policy as I have seen during the last three years.”

The implementation of GILS took place during a period of significant technological and agency change, uncertainty, political discord, opportunity, pressure, stress, and excitement for Federal information managers. *One important finding from the study is that GILS, given this context, simply was unable to compete for the attention, resources, and commitment from most agency administrators.*

4.2. EXTENT OF CURRENT GILS IMPLEMENTATION

An initial analysis of the number of agencies involved with GILS implementations presents a relatively positive picture. During the evaluation, the investigators identified 45 units of government (e.g., executive agencies, independent agencies, commissions, government corporations, etc.) carrying out some type of agency GILS implementation. A closer look, however, reveals the extent to which these agencies have taken ownership of the initiative. Further, certain cabinet-level departments appear not to have undertaken any GILS implementation as of February 1997 (e.g., Departments of Education, Justice, Transportation, and Veterans Affairs).

Agencies had the option of either mounting their GILS records on an agency server that complied with *FIPS Pub. 192* specifications for using Z39.50 or contracting with another agency to make their records available. The study identifies these approaches as “record-source hosted GILS” and “brokered-GILS.” The former means that the agency creating the records is also responsible for making those records available via the Internet, and the latter means that an agency creating GILS records contracted with another agency to make those records available. Both the Government Printing Office (GPO) and FedWorld offer this “brokering” service to agencies.

The study identified eighteen “record-source hosted GILS” sites where an individual agency server provides access to that agency’s GILS records. A total of 2,089 GILS records are available from these servers. Table 4-2 presents a summary of records provided by each agency. See Appendix B for a list of agency GILS servers/databases with network addresses.

FedWorld and GPO offer services to agencies in mounting and making agency GILS records accessible, and through this service they have become central points of access to the majority of agency GILS records. Table 4-3 summarizes the brokered records from GPO and FedWorld.

As of March 1997, GPO hosted a total of 2,815 GILS records from 27 agencies (in addition to mounting the Privacy Act notices database from NARA). It also provides “pointer records” to 7 agencies that have GILS records available but which are not mounted at GPO. In April 1997, GPO began offering a new search application through which a user can submit a search across one or more agency GILS, whether or not the records are mounted at GPO. A user selects which agency GILS databases or servers to search, submits a query, and the search is broadcast to the selected GILS databases and servers. GPO’s recent efforts point to one direction of possible cross-agency, government-wide searching with GILS.

GPO also offers searches on a database compilation of Privacy Act Notices, an area of GILS coverage mandated by OMB Bulletin 95-01. In August 1995, NARA and OMB agreed that this requirement could

be met by allowing NARA to make available its compilation of *Federal Register* Privacy Act notices on GPO (see Appendix A-4 for the NARA memo). This agreement relieved agencies from the requirement to create GILS records for agency Privacy Act systems and associated notices already published in the *Federal Register*. The NARA database of *Federal Register* notices provides coverage of additional agency resources not necessarily reflected in the records in Tables 4-2 and 4-3. GPO mounted the compilation of Privacy Act Notices to meet the requirements of OMB Bulletin 95-01. There are currently 5483 documents listed in the Privacy Act Notices compilation, but these are not in the standardized GILS record structure, and are not calculated into the total number of GILS records available for searching by users.

Also, as of March 1997, FedWorld served as host for three agencies' records totaling 353. During the course of the evaluation study, however, FedWorld expanded its listing of GILS records to include those hosted by GPO (excluding the GPO Privacy Act application) and an additional six record-source hosted GILS sites. FedWorld currently offers searches of 35 different agencies' records. (FedWorld lists 36 agencies' databases but that number includes Department of Commerce GILS records mounted at FedWorld as well as its records mounted at GPO.) Users of FedWorld GILS can search the three databases mounted at FedWorld as well as following links to other agencies' GILS records.

Table 4-2
Record-Source Hosted GILS

Record-Source Hosted GILS	Total Records	Source Date*
1. Department of the Interior	322	3/19/96
2. Department of Agriculture	135	3/6/97
3. Department of Defense	494	3/6/97
4. Department of Energy**	6	Not available
5. Department of Labor	34	Not available
6. Environmental Protection Agency	239	3/6/97
7. General Services Administration	46	12/29/95
8. Health and Human Services	642	2/13/97
9. Housing and Urban Development	5	Not available
10. National Aeronautics and Space Administration	11	1/5/96
11. National Archives and Record Administration	37	3/6/97
12. National Labor Relations Board	7	Not available
13. National Transportation Safety Board	5	Not available
14. Office of Management and Budget**	3	Not available
15. Small Business Administration**	39	2/4/97
16. Tennessee Valley Authority	3	3/1/96
17. United States Postal Service**	15	11/15/95
18. Department of Veterans Affairs	46	3/21/96
(Minimum = 3; Maximum = 642; Average = 116)	TOTAL	2,089

* Date associated with the number of records found; these sources were checked in March 1997

** GILS records offered as standalone HTML files rather than in a WAIS or Z39.50 searchable/accessible database

BEST COPY AVAILABLE

Table 4-3
Brokered GILS

GPO and FedWorld Brokered GILS	Total Records	Source Date*
GPO-Brokered GILS		
1. Consumer Product Safety Commission	34	6/18/96
2. Department of Commerce	281	11/5/96
3. Department of State	95	6/18/96
4. Department of Treasury	594	12/26/96
5. Equal Employment Opportunity Commission	26	6/18/96
6. Farm Credit Administration	5	6/18/96
7. Federal Communications Commission	39	6/18/96
8. Federal Emergency Management Agency	4	6/18/96
9. Federal Labor Relations Authority	9	6/18/96
10. Federal Maritime Commission	12	8/14/96
11. Federal Reserve Board	1	6/18/96
12. Federal Trade Commission	10	6/18/96
13. General Services Administration	2	2/4/97
14. Government Printing Office	36	3/3/97
15. International Trade Commission	11	7/30/96
16. Merit Systems Protection Board	8	6/18/96
17. Office of Government Ethics	11	6/18/96
18. Office of Management and Budget	3	6/18/96
19. Office of Personnel Management	15	6/18/96
20. Overseas Private Investment Corporation	9	6/18/96
21. Pension Benefit Guaranty Corporation	17	11/26/96
22. Railroad Retirement Board	13	8/28/96
23. Securities and Exchange Commission	139	10/18/96
24. Selective Service System	9	6/18/96
25. Social Security Administration	1,203	6/18/96
26. U.S. Commission on Civil Rights	223	6/18/96
27. Nuclear Waste Technical Review Board	6	6/18/96
TOTAL	2,815	
(Minimum = 1; Maximum = 1,203; Average = 104)		
Privacy Act Notices compilation at GPO	5,483	
 FedWorld-Brokered GILS		
1. Federal Energy Regulatory Commission	14	2/14/97
2. U.S. Nuclear Regulatory Commission	48	2/14/97
3. Department of Commerce	291	2/14/97
TOTAL	353	
(Minimum = 14; Maximum = 291; Average = 118)		

* Date associated with the number of records found; these sources were checked in March 1997

Based on the information presented in Tables 4-2 and 4-3, a reliable estimate of the number of available GILS records (as of March 1997) is approximately 5,000. One might immediately ask:

Is the 5,000-plus GILS records that have been created an appropriate number of records (either in total or per agency) for carrying out the mandate of GILS?

This, however, is a difficult question to answer. The Information Infrastructure Task Force (IITF) report (Information Infrastructure Task Force, 1994, p. 11), provided the following estimate:

The entire GILS Core is not likely to contain more than 100,000 locator records. In addition to locator records for information systems, it is estimated that the GILS Core will contain up to 1,000 locator records for each Federal agency that is a major disseminator of public information. Agencies that are not major disseminators will typically have fewer records in their portion of the GILS Core, especially if the agency is relatively small.

Although the origin of the "100,000" number is unclear, the goal was to create sufficient GILS records to provide comprehensive coverage of Federal government information resources and assist users in locating those resources. The estimated 100,000 locator records would describe the resources identified in OMB Bulletin 95-01: automated information systems; locators to agency resources; and Privacy Act Systems.

OMB 95-01 defines GILS Core as "a subset of all GILS locator records which describe information resources maintained by Federal agencies, comply with the GILS core elements defined in *Federal Information Processing Standards Publication (FIPS Pub.) 192*, and are mutually accessible through interconnected electronic network facilities." The OMB 95-01 definition is less descriptive than that offered in the IITF report which provides additional information about GILS Core including:

The GILS Core will include records for all information locators that catalog other publicly accessible information resources at least partially funded by the Federal government, as well as for each of the Federal government information systems that include publicly accessible data or information. While GILS Core records can point to any kind of information source, they are especially designed for helping users navigate among a wide array of other locators in various formats. It is not recommended that agencies use the

precise format of the GILS Core locator records to describe all types of information resources.

The emphasis in the GILS design document and policy on distinguishing "GILS Core" records from other GILS records, however, has not led to clear distinction in practice. The analysis of a sample of GILS records conducted as part of this study (see Appendix E-2) showed little difference between GILS records identified as "Core" (through the use of the term "U.S. Federal GILS" in the Controlled Vocabulary-Local Subject Index Term element) and those not so described.

If an agency already had locators or inventories that could be described by a GILS record, a few GILS records might be sufficient to address the goal. In the absence of pre-existing locators, however, some agencies have been describing individual documents and publications. In that case, a major information disseminating agency might have to create thousands of records to gain the coverage envisioned for GILS. Measuring the extent of coverage of agency resources by GILS records would require the existence of comprehensive inventories of agency resources (i.e., a baseline against which to measure). Although the study did not attempt such a measure of coverage, the question of whether 5,000 records is sufficient to provide users with the ability to discover and access agency resources needs to be addressed. The question can be framed as follows:

Are we moving towards government-wide coverage of publicly available government information through the GILS records?

Data from the study suggest that the GILS initiative, as it is currently being carried out, is not likely to improve coverage. Moreover, users, who were the focus of this evaluation, stated that based on their experience with GILS, current coverage of government information resources is insufficient. Users also want GILS to provide direct access to the actual information resources.

One can claim that the GILS initiative is new, and the approximately 5,000 records created in the past 2 years are a good beginning. Other findings discussed below, however, suggest that many agencies are not

likely to be creating new records. Thus, the current 5,000 may be the extent of GILS record creation and government-wide information resource coverage cannot be expected. For example, one major information disseminating agency stated that their approximately 300 records cover what GILS mandated, and it is not likely to be creating additional GILS records. A number of the smaller agencies stated that because insufficient resources were allocated to implement GILS, and because they see little return on investment (ROI), they would not be creating more records and in fact would not maintain the records they had created. The Source Date column in Tables 4-2 and 4-3 is indicative of GILS activities, with many of the databases showing the most recent updating in 1996.

To estimate the universe of GILS records, it was first necessary to identify existing GILS sites. This was a major task to ensure that no agency involved in any GILS implementation was overlooked. Reviewing the steps in that identification process (see below) also demonstrates one of the challenges facing users of GILS and a liability of the current implementation—there is not a single registry of existing agency GILS implementations.

Implementing such a registry would provide a user with a source to determine which agencies have GILS implementations, the number of records associated with each implementation, and the network location of each implementation.

For the evaluation study, the sites listed in Tables 4-2 and 4-3 were discovered through the following activities:

- Verbal or written mention during the 1996 GILS Conference presentations and in handouts and survey responses
- Linking from the White House Web site's "President's Cabinet" (http://www.whitehouse.gov/WH/Cabinet/html/cabinet_links-plain.html) and "Federal Agencies and Commissions" (http://www.whitehouse.gov/WH/Independent_Agencies/html/independent_links-plain.html) to agency homepages, which in turn linked in some cases to FedWorld GILS (<http://fedworld.gov/gils>)

- Web searches by means of Alta Vista and Lycos search engines for Executive department and agency names
 - as delineated in the 1996-97 *Government Manual* via GPO Access (http://www.access.gpo.gov/su_docs/aces/aaces002.html)
 - as comprising the Chief Information Officer Council as specified in Executive Order 13011 of July 16, 1996 *Federal Information Technology* (<http://www.gsa.gov/irms/ka/regs/exo13011/exo13011.htm>)
- WEB searches by means of Alta Vista and Lycos search engines for "GILS," and "government information locator service"
- GPO Access GILS server
- Appendix A of *Potholes on the Information Bridge to the 21st Century*, the Second Annual OMB Watch report on the U.S. Federal Government Information Locator Service (Henderson, 1997).

This effort was necessary to ensure that all agency GILS sites were identified. Through this effort, the investigators not only confirmed the agencies' GILS identified at GPO and FedWorld, but also identified 8 other agency GILS implementations not listed by either GPO or FedWorld. Not all of those 8, however, have their GILS records residing on an information retrieval-based platform such as WAIS or Z39.50-compliant server. These agencies offer their GILS records via a Web server, and the GILS records are simply hypertext markup language (HTML) files comprising GILS elements. The fact remains that these agencies are implementing GILS in a fashion, and their records should be included in estimating the universe of GILS records.

FedWorld's and GPO's recent efforts to provide single points of access to multiple agencies' GILS records move the Federal GILS initiative in the direction of a truly government-wide locator service. The study found, however, a range of responses to and interpretations of what GILS is or should be and how it should be implemented. These responses and interpretations by individual agencies may mitigate against comprehensive coverage of publicly available government information in a manner that is useful to

the public and other users trying to discover and access government information. The first "opportunity" that needs to be addressed is how to refocus the GILS effort by clarifying its purposes, goals, benefits, and expected impacts.

4.3. OPPORTUNITY: REFOCUS GILS FOR CLARITY OF PURPOSE AND UTILITY

Many of the findings reported in this section reflect a need for a clarification of what GILS is, what functions it should support, what agencies are

expected to do, and what benefits might accrue. The study found that the original expectations for agency participation in GILS did not adequately acknowledge the resulting burdens upon many agencies nor account for a range of factors that might constrain agency GILS implementations (e.g., the lack of appropriate network and information technology infrastructure). On the basis of these findings, the investigators recommend refocusing the U.S. Federal GILS efforts in the next stage of GILS development. Table 4-4 summarizes the findings and recommendations for this opportunity.

Table 4-4
Refocus GILS for Clarity of Purpose and Utility

OPPORTUNITY: REFOCUS GILS FOR CLARITY OF PURPOSE AND UTILITY	
Findings	Sources of Evidence*
4.3.1. People Are Confused about GILS Mission, Purposes, and Uses	CA, FG, KP, SU, SV, US
4.3.2. Expectations for GILS Are Evolving	FG, SU, SV
4.3.3. Government-Wide Administrative Coordination and Policy Oversight Are Lacking	FG, KP, SU, SV
4.3.4. Smaller Agencies Feel Special Burden and Frustration	FG
4.3.5. Agencies' Cultures and Missions Promote Different Commitment to GILS	FG, KP, SV
4.3.6. Intra-Agency Efforts Reflect Different Levels of Enthusiasm for GILS	FG, SV
4.3.7. GILS Benefits Compared to Burdens Are Not Clear	FG, KP, SV
Recommendations	
4.3.8. Focus on Public Access to Government Information	
4.3.9. Focus Scope of Descriptions On Network-Accessible Information Resources	
4.3.10. Identify Responsibilities and Authority for Policy Leadership, Government-Wide Coordination, and Oversight	
4.3.11. Implement a Refocused GILS Initiative	
4.3.12. Require Agency Reporting on GILS Progress and Reward Agencies That Achieve Stated Objectives	
4.3.13. Ensure Ongoing, User-Based Evaluation for Continuous Improvement	

* CA=content analysis of GILS records; FG=focus group sessions; KP=interviews with key participants; LA=log analyses of Web servers; SU=survey conducted at the 1996 GILS Conference; SV=site visits to selected agencies; US=scripted online user assessments of GILS

4.3.1. FINDING: People Are Confused About GILS Mission, Purposes, and Uses

Considerable confusion exists among both agency implementors and external users as to the purpose of GILS, what it was intended to accomplish, and just "what exactly the GILS is." One person commented

"at 30,000 feet, GILS is a good idea, but implementing this at ground level, it became all things to all people." This problem is exacerbated by different stakeholder groups and audiences who each look at the GILS initiative from different perspectives. Figure 4-1 summarizes some of the competing purposes and audiences that *might* be

addressed by GILS. Clearly, additional possible purposes and audiences could be added to this figure.

The confusion over what GILS was intended to be, what it is, and what it might become was a constant theme in the various data collection efforts. As one example, the survey administered at the November 1996 GILS conference asked several questions related to respondents' understanding and definitions of GILS. Approximately 180 conference participants completed the survey (see Appendix E-1 for details on survey respondent demographics).

Question 1 asked respondents for their definition of GILS. This open-ended question produced a wide range of answers. (Tables AE1-7 through E1-10 in Appendix E-1 summarize the responses.)

Respondents' definitions highlighted four primary perspectives on GILS, but their definitions often addressed more than one:

- GILS from the perspective of functions including Finding Aid ("card catalog," "index," "pointers," etc.); Access ("provide access to," "retrieve information," etc.); IRM ("managing resources," "records management," etc.); Collect ("agencies 'collect' information via GILS"); Control ("agencies 'control' information via GILS")

- The types of information GILS comprises including Publications, Resources, Systems, Records, and Services
- Potential users of GILS including Public, Agency, Private, Library, Researchers, etc.
- The coverage of GILS including "Federal government information," "Important/major/prime information," "Executive information," "Electronic information," "Useful information," and "Other." The category of "Other" includes the following limitations to GILS coverage:
 - A basic replacement and improvement to requesting information from Pueblo, CO—you can find all agencies with information on topic
 - [primary] systems of records
 - Certain federal holdings
 - Information federal agencies choose to make available
 - Government services policy procedures information
 - Public records to patrons of the service
 - Records federal agencies are creating
 - Technical knowledge gained through research
 - All of IRS systems
 - Information for government agencies to complete daily duties.

Figure 4-1
Clarifying GILS Purposes and GILS Users

Possible Purposes	Possible Users (among many...)				
	Records Manager	Librarian	'The Public'	FOIA Officer	Program Manager
Create locators to government metadata					
Identify specific government information or records					
Access FOIA information and records					
List major information systems					
Inventory Privacy Act Notices and systems					
Cross-agency search/retrieval of metadata (or information)					
Provide links to GILS in states and other countries					
Identify Federal records that need to be scheduled					
Provide records retention schedules					

The broad range of responses to this request for a definition of GILS is indicative of competing expectations as well as misconceptions on the part of users and implementors.

The survey also requested respondents to assess a number of key issues, some of which addressed definitions and purposes of GILS (see Tables E1-11 and E1-12 in Appendix E-1). There was a high level of agreement to the statement: *A purpose of GILS is to improve public access to government information* (89% of respondents agreed with this statement). Yet only 55% agreed with the statement: *A purpose of GILS is to help agency officials better manage agency information*. Only 45% of respondents agreed with the statement: *I am able to describe GILS accurately and fully to others*. In terms of coverage of GILS, only 33% agreed with the statement: *GILS records represent the complete information resources of an agency*.

The site visits and focus groups also highlighted a lack of clarity about the purpose of GILS. Many of the participants in those activities identified the need to clarify the purpose of GILS so that people (e.g., agency staff and public users) could know what to expect to find when using it. One person in an agency site visit stated that "GILS has an identity crisis—what exactly is its purpose? Is it for public relations? Is it for providing information to the public? Is it for records management?" The need to clarify GILS' purposes and objectives was also tied to understanding what tangible benefits would accrue to agencies by using GILS.

To a large extent, GILS has become "different things to different people" or, more precisely, people see in GILS what they want to see. Individuals complained that they cannot find quick factual answers to reference questions in GILS. While users might have such exceptions, the fact is that the original design of GILS did not intend it to support that functionality. Others have proposed that GILS be used to manage electronic Freedom of Information (EFOIA) requests and information—again, never a stated goal or purpose of GILS. In both of these instances, the GILS record structure does not support such purposes.

The study found contradictory, confused, ambiguous, and erroneous perceptions of GILS' intended purposes and GILS' potential purposes. The investigators were told of instances when GILS policymakers and implementors, during early training sessions, publicly disagreed with each other as to GILS' purposes.

Given this situation, the successful implementations were those by agencies that decided for themselves what GILS would be in *their* setting. For example, EPA, Defense, and Treasury created agency GILS to serve both internal and external users and uses. EPA sees its GILS implementation as a component of its larger information dissemination and access responsibilities. Defense and Treasury see GILS as serving as a useful tool for inventorying and information management. While these are not contradictory roles for GILS, a user looking for information across the government may be confused by the differing levels of coverage, granularity of description, and focus of specific agency GILS.

Study participants and users of GILS judge the service in light of *their* perceived purposes and expectations of GILS and often are very disappointed. Clearly, some of the cells in Figure 4-1 are not mutually exclusive. But the findings indicate that there is a lack of agreement as to the purposes of GILS and what one can reasonably expect GILS to accomplish in terms of providing access to and management of government information. One person commented that GILS does not provide *government-wide information* (as advertised in the name of the service); rather it identifies some possible *agency* sources that *might* have the information needed if one *could* get into those other sources. To this person, the name of GILS was a misnomer in itself.

4.3.2. FINDING: Expectations for GILS Are Evolving

At the 1996 GILS Conference, a number of speakers made an important point by separating the original GILS vision from the manner in which agencies had implemented GILS to date. The GILS Conference survey (and presentations made at the Conference) and other data collection activities indicate substantial support for the original GILS concept of improved public access to government information. Yet only

limited support exists for the GILS implementation as outlined in OMB Bulletin 95-01 or as undertaken by most agencies. This may be due, in part, because a "government-wide" perspective on Federal information has yet to emerge from the GILS initiative.

Study participants noted the desire to obtain the "actual" information rather than simply descriptions of information resources. In part, the widespread deployment of Web technology has raised expectations on the part of users in terms of gaining immediate full-text access to government information.

This study found support for what might be termed a *refocused* GILS concept which can be summarized as:

An easy-to-use and coherent government-wide information search service available from one or more service points that enables users to discover, locate, select, and access publicly available government information resources (e.g., agency information systems, specific information dissemination products, and existing locators to those products) through standardized metadata that describe those resources and provide direct links to the described resource (e.g., full-text documents, other online services).

Study participants suggested the original GILS concept is being replaced by a belief that a refocused GILS is of greater utility. This refocused GILS concept is not incompatible with the existing concept of GILS, yet it is more limited in scope (e.g., the refocused GILS is not tied to records management; see Section 4.5.1.). In addition, the refocused GILS clearly responds to the desire of users for a single point of access for searching government-wide for information. This can be seen as a positive evolution for GILS.

As noted in Chapter 1, GILS was an ambitious undertaking. The effort should not have been viewed as a panacea for the various issues relating to access and management of government information, and it could be expected that major technical and policy issues would arise during this

early implementation period. Learning from implementation experience has been common for many agencies. Further, the technology environment in which GILS has been implemented since early 1995 has changed enormously. The emergence of Web technology has generated new expectations among Internet users, and a simple set of pointers to metadata is no longer sufficient for most users.

These and other factors have created a need for a more focused and consensus-driven conception of GILS that responds to the demands of users, both Federal agency staff and non-government users, interested in discovering what information is available and then being able to access that information directly.

4.3.3. FINDING: Government-Wide Administrative Coordination and Policy Oversight Are Lacking

GILS, as originally conceived, would be a decentralized information service consisting of agency information locators linked and interoperable through the use of common technical and content standards. OMB Bulletin 95-01 identified lead agencies for particular aspects of GILS (e.g., NARA for record creation guidelines and training). The Bulletin, however, was silent on how government-wide coordination and oversight of the GILS initiative would occur. The Bulletin established the GILS Board with responsibilities to "evaluate the development and operation of the GILS," but it has met only once since the publication of OMB Bulletin 95-01. Study participants suggested that a lack of government-wide coordination and oversight is one of the causes for the current state of GILS. Further, a number of participants recognized that the decentralized implementation of GILS needs to be balanced by some level of centralized management and coordination to assure the coherent development of a government-wide information locator service.

One group that has been active since March 1995 is the Special Interest Group on the Government Information Locator Service (GILS SIG). Its Statement of Purpose (see Appendix A-6 for the complete Statement) includes the following:

The purpose of the GILS Subgroup is to help fully realize the potential of the Government Information Locator Service (GILS) concept, and to promote the development and use of this open systems approach for information search and retrieval. The Subgroup exists to help organizations implement GILS, and also to encourage effective evolution of the GILS standard to meet new uses. To accomplish these purposes, the GILS Subgroup: 1) serves as an open forum for the exchange of ideas on GILS development, use, and refinement, 2) forwards to the OIW/SIG-LA appropriate recommendations for changes to GILS, and 3) promotes sound implementation and broad public awareness of GILS. One emphasis of the Subgroup is to strengthen the U.S. Federal GILS to provide a model and test case for other GILS implementations.

The GILS SIG has been instrumental during the past 2 years of GILS implementation and has provided a forum for information sharing during GILS development. It is not authorized, however, as a policy making or coordinating body for U.S. Federal implementations of the GILS Profile. Since the GILS Profile has application outside of the U.S. Federal implementation, the GILS SIG membership is open to anyone interested in using the GILS Profile (e.g., state and other national governments). Since its responsibilities and participants are broader than U.S. Federal implementation of GILS, it is not an appropriate forum for administrative and policy coordination for the U.S. Federal GILS initiative.

The GILS SIG operates under the auspices of the Open Systems Implementors Environment Workshop (OIW) and assumed in late 1996 the responsibility for maintaining the GILS Profile. The GILS SIG does not provide a formally constituted or authorized forum for discussions of U.S. Federal implementations of GILS. As originally constituted, the OIW groups were places where implementors and users could convene to identify specific application requirements for standards and to arrive at consensus agreements on profiles. Given this, U.S. Federal implementors of GILS are just one

user group that would bring their requirements to the GILS SIG (along with Canadians, states, etc.).

Generally, no administrative unit has provided government-wide leadership, coordination, and development for the Federal GILS. GILS "leadership" that does exist has occurred at the agency level and resulted because of strong administrative interest and commitment by the individual agency (e.g., EPA and Defense).

The 1996 GILS Conference survey asked respondents for their assessment of the following statement: *There is adequate policy guidance to direct the development and operation of GILS.* Only 39% of the respondents agreed with this statement. To be fair, however, one should note that less than half of the respondents were familiar with OMB Bulletin 95-01 and other GILS documents and policies (see Table 4-10 below). Overall, study participants generally agreed that the existing GILS policy provided too much latitude to agencies, that OMB, Office of Information and Regulatory Affairs (OMB-OIRA) had "shirked its duty" to enforce GILS provisions, that OMB-OIRA provided conflicting messages to agencies about the relative importance of GILS development, that agencies rarely had an internal policy on GILS development and management, and that with the significant amount of information policy issues that have been on the government's agenda during the past 2 to 3 years, GILS policy and oversight fell through the cracks.

Spokespersons for various agencies—small and large—believe that after OMB finished Bulletin 95-01 it simply "dropped the ball" in terms of administrative leadership and policy oversight. Others, however, believed that such administrative leadership and oversight were not the responsibility of OMB. Whatever one's point of view, the study found that the lack of administrative leadership and coordination of GILS implementation across agencies and the lack of oversight to determine the degree to which agencies were in fact complying with OMB Bulletin 95-01 contributed to the current limited success of the GILS effort. Centralized leadership, coordination, management, and oversight is critical as a counterweight to the decentralized, distributed implementation of GILS as a networked service.

4.3.4. FINDING: Smaller Agencies Feel Special Burden and Frustration

Participants in the study from small agencies felt burdened and isolated, and believed they were not heard regarding GILS. Smaller agencies were especially frustrated with the lack of leadership, direction, and resources during the GILS implementation process. They expressed significant dissatisfaction with OMB. In particular, they felt OMB had not listened to or acknowledged the burden that GILS would impose on their agencies. Individuals at these agencies translated reinventing government as "doing more with less," and, with GILS, it was doing something more with questionable value. They felt disenfranchised from the process of developing GILS, and viewed GILS as something directed primarily at the larger agencies—those that had the resources to implement GILS. While the larger agencies may hold the bulk of government information, GILS, if it is to be a government-wide information service, must have government-wide coverage. From this perspective, smaller agencies have many important information resources to contribute.

GILS implementors in many small agencies have responsibilities not only for records management but also computer security, FOIA, etc. The requirement to implement GILS in addition to these other responsibilities seemed unreasonable and many were quite angry about having to manage such a range of responsibilities. Thus, a number of these agencies are barely, if at all, carrying out the directives that govern GILS. While they have created *some* GILS records, and those records are accessible (usually on a brokered basis by GPO), a number of participants indicated no plans to produce additional records or maintain the records they have created. OMB Bulletin 95-01 required that agencies must create locator records, so some records were created, period.

For many of the smaller agencies, inadequate technology infrastructure or technology resources was a constraining force in accomplishing the GILS mandate. But such infrastructure constraints are not necessarily limited to the smaller agencies. At least

one of the larger agencies visited by the investigators described the lack of a robust networked infrastructure (e.g., lack of network access at the desktop by those creating GILS records) and its impacts on implementing GILS. For example, the use of distributed data input procedures and software such as that developed by the Defense Technical Information Center (DTIC) was not an option if a modern information technology and network infrastructure did not exist in the agency.

Many of the smaller agencies did not believe mounting their few GILS records on a local agency Z39.50 server was cost effective. As an example, more than 20 of the smaller agencies contracted with GPO to mount their records in the interest of resource optimization. Yet this expediency resulted in a quality-control constraint; agency staff that lacked desktop network connection could not access the records once they were sent to GPO, and thus could not update records easily.

Based on discussions with representatives from small agencies, the investigators found that as a group, the small agencies are unlikely to participate in future GILS activities without significant changes in the existing GILS initiative. Their participation will be contingent upon the degree to which they are involved in future GILS planning, the degree to which they better understand GILS initiatives and benefits, and the degree to which they can marshal resources to be compliant with requirements. The latter will require some demonstration of tangible benefits (and the costs incurred for those benefits) of extending their GILS implementations.

4.3.5. FINDING: Agencies' Cultures and Missions Promote Different Commitment to GILS

Where an agency has a history of strongly supporting public access to its information resources, GILS tends to be more enthusiastically embraced and perceived as successful than in agencies without such a history. Where top management has endorsed GILS and provided strong support—especially by dedicating staff and capital—GILS has tended to be much more successful, at least in its implementation if not in its use. Shallow administrative support, no agency champion, and convenience-based decision making

(e.g., choosing a GILS record data input/creation software because it was virtually free) severely constrained GILS success. As a corollary, when staff *asked* to be in charge of their agency's GILS effort or were already committed to the GILS concept, the agency's GILS efforts were more likely to be a success.

Some agencies already had some type of a locator or finding tool in place. In these agencies (e.g., EPA, Defense, and GPO) the GILS effort appeared to be better understood and coincided with existing agency culture that was predisposed to support public access. A number of other agencies did not have a culture predisposed to support public access. A participant in a focus group with representatives from Federal agencies declared that except for one or two items, her agency's information resources contained proprietary or private information that would not be made public; she questioned why she should create GILS records identifying those resources.

Champions who were dedicated to the GILS concept, knowledgeable about locators and public access, and had good credibility in the agency were critical factors in successful agency implementations. One or two competent staff working at the day-to-day level, providing continuous injections of enthusiasm, and helping to solve problems can, and did, make the difference between a successful and unsuccessful effort. The study found only a limited number of agencies where the existing culture, administrative support, and the involvement of champions directly supported the GILS effort.

OMB Bulletin 95-01 delegated primary responsibility for implementing GILS to the departments and agencies, who then had considerable freedom to determine how they would respond. The findings identify three basic types of agency response to the GILS initiative:

- **Thoughtful and Committed:** A small number of agencies carefully planned their agency response to the GILS initiative, had a champion, provided staff and other

resources to support the effort, and produced a working GILS.

- **Good Faith Effort:** In these agencies, someone or some unit emerged to motivate production of at least some GILS records and meet "the letter of the law" even without agency-wide support or commitment to the GILS concept or its implementation.
- **Minimal Compliance:** For a number of agencies, there was little to no acknowledgment of GILS. These agencies followed the letter of the law (in their interpretation), and did so by producing a handful of records—usually mounted by a brokering agency—and then considered their GILS effort completed.

These three types characterize those agencies providing some GILS product. It should be noted, however, there are some departments and agencies that have yet to engage in any GILS development (e.g., Departments of Education, Justice, Transportation, and Veterans Affairs).

Given this wide range of responses, generalizations of the findings from an agency perspective are difficult to make. Indeed, it should be kept in mind that there are a number of different agency-based GILS and *not* one GILS.

4.3.6. FINDING: Intra-Agency Efforts Reflect Different Levels of Enthusiasm for GILS

Staff responsible for implementing GILS quickly came up against the reality that different agency units had different levels of enthusiasm for GILS. Some individuals who were tasked to "handle" the GILS initiative in their department or agency found the job to be very onerous, especially since the task came without additional resources. Others latched onto the task and were extremely enthusiastic about the GILS initiative as a means to improve access to government information, or for realizing other individual or agency-specific benefits (e.g., the individual had a personal commitment to GILS or GILS was viewed as a useful information management tool).

Cooperation among staff within departments and agencies tended to vary with individual agency units' perception of the importance of GILS. Some factors unrelated to GILS worked in its favor, such as when an agency suffering from a negative public image seized upon GILS as a way to improve its image by providing access to information about the agency. In most instances, persons assigned responsibility for GILS had little direct authority over others from whom the person had to obtain records information. Participants in this position reported no enthusiasm for the GILS effort and, in some instances, outright anger about "having to do this on top of everything else that I am supposed to do."

Site visit and focus group participants identified one barrier to implementing GILS as the difficulty in obtaining agency-wide staff involvement in gathering information to create GILS records. Agencies that had preexisting information locator resources found this part somewhat easier because they had already established procedures for locator data collection and input. Most agencies believed that responsibility for GILS records input should reside with the "offices of primary interest" (i.e., the office or staff responsible for a particular information resource) but obtaining these offices' cooperation was a chronic problem. Often personnel in these offices saw GILS records input as just one more work demand. In some cases, these staff resisted GILS because they believed that putting their names and phone numbers into GILS records as contact persons would increase their workload.

The study finds a significant likelihood that (1) some "minimal compliance" agencies will not create many additional records nor update ones originally submitted, and (2) those agencies that are conscientious about their GILS efforts will find it increasingly difficult to obtain updated information. These findings point to a possibility of overall GILS degradation over time.

4.3.7. FINDING: GILS Benefits Compared to Burdens Are Not Clear

The study found a range of views on the benefits versus the burdens of GILS. Many agency personnel see GILS as a pure burden without benefit. Or worse, they see it as an unfunded mandate for which they had no administrative commitment or resources, and which distracted them from other "more important tasks." In site visits and focus groups, emotions often ran high reflecting the anger and frustration felt by some agency implementors. They were on the receiving end of the mandate to implement GILS and concluded that the entire effort was a waste of time and effort, without regard to obtaining additional resources. They were quick to point out that they believed strongly in improved public access to government information. But, in their view, GILS, as currently conceived, "was certainly not the tool to accomplish improved access, nor did it assist in records management efforts." A number of these respondents argued that GILS was "dead on arrival."

Another group of respondents thought GILS will return little benefit if it remains an isolated system. This view holds that GILS becomes useful only when integrated into other systems such as agency Web sites, other information systems, or other metadata schemes. Many questioned whether existing levels of GILS use and benefits warrant continued support and development. Others were unable to articulate any specific tangible benefits arising from GILS. On the other hand, these same people often tended to believe GILS should not be eliminated, but rather refocused and improved.

Yet a final group of agency implementors had a much more positive assessment of GILS and listed a range of specific benefits that had accrued to their agency as a result of their GILS implementation efforts. Benefits mentioned include:

- Improved public access to electronic and other agency information resources
- Improved agency knowledge and coordination of existing information resources and how to access them
- Better understanding of the importance of metadata and the need for metadata records

- Increased visibility and involvement for the IRM, information managers, records managers, etc. in department/agency/bureau information resources management
- Identification of potential resources that may need to be scheduled for records retention and preservation
- Development of GILS as a "platform" or base from which other systems could be linked into a "one stop shopping" approach for locating and accessing government information.

This group provided the investigators with a number of anecdotes and experiences that supported these benefits.

Participants of several focus groups believed that GILS is, in fact, serving as a catalyst for "good things" that should get done in the area of information access. One benefit people pointed to was the fact that, as a result of GILS, agencies were indeed taking inventories of their information products, which is something they were expected to do but often had not. On the other hand, some voiced the fear that GILS is "robbing resources from other information access efforts that are more worthwhile."

To some degree, GILS burdens and benefits are in the eye of the beholder. There also was a clear correlation between those agencies that had committed staff, resources, and administrative support to also believing they had gained significant benefits from the effort—the opposite correlation also holding true as well. Since no formal cost-benefit study has been done on the GILS effort, and was not completed as part of the current study, the study finds that perceived benefits are likely to be situational and stakeholder group dependent.

Notwithstanding the varying purposes and goals discussed earlier in this section, GILS was premised on improving public access to government information, agency information management, and records management. Another way to think about GILS is: what is an appropriate and realistic purpose for GILS whereby it provides tangible

benefits to agency implementors and provides a value to users who want to discover, identify, and access government information?

4.3.8. RECOMMENDATION: Focus on Public Access to Government Information

Early in the evaluation study, it became apparent that "GILS" meant different things to different people. While there was some consensus that the U.S. Federal GILS initiative was intended to support and enhance access to government information, there was little consensus on exactly how that would be accomplished. GILS policy statements and implementation goals raised high and varied expectations of GILS. Unrealistic expectations of what GILS could accomplish has in part increased the volume of expressed disappointment and frustrations by both agency staff and users.

That GILS has been many things to many people is no accident. OMB Bulletin 95-01 identifies several purposes and goals for GILS:

- Assist users in locating government information by developing core locator records for
 - Information dissemination products
 - Automated information systems
 - Privacy Act record systems
- Scheduling and disposition of records through NARA
- Electronic records management
- Improved agency responses to the Freedom of Information Act (FOIA) requests
- Potential reduction of information collection burden on the public.

The question is: can one mechanism such as GILS serve multiple and diverse purposes and goals?

On the basis of policy goals for GILS as well as what the investigators learned in the study, it is possible to identify purposes that stakeholder groups have assigned to or expected of GILS including:

- Inventorying of selected agency information resources

- Capturing and creating metadata for those resources
- Making the metadata available for public access
- Using the metadata for records management
- Linking metadata records to actual full-text information resources
- Enhancing public knowledge of and access to government information
- Providing full-text access to government information.

Because confusion exists among the agencies and the public as to what GILS is and why it needs to exist, the investigators recommend that the Federal GILS initiative be refocused and part of the process of refocusing should be a redefinition and clarification of the purpose and goals of GILS. In addition to clarifying the purpose, scope, and expected functionality of GILS, the task of refocusing should address a range of questions such as:

- What demonstrable benefits result from implementing GILS?
- What strategies are appropriate for marketing the GILS "product" to agencies and users?
- What types of training are required to accomplish GILS objectives?
- How can agencies cooperate to develop one-stop shopping by subject?
- On what basis should agencies establish electronic linkages between GILS and full-text information resources and electronic services?
- What is a desirable level of granularity or units of information described by GILS records?

The experience to date with GILS (as a technology implementation as well as an information policy initiative) suggests that loading any one system with too many expectations reduces the likelihood that it can adequately fulfill any of the expectations.

The investigators recommend that the primary purpose of a refocused GILS initiative should be to assist users in the discovery, identification, and access of government information (in the broader networked environment this is referred to as networked information discovery and retrieval). The investigators heard from many people in the study that an information locator service should assist people in finding out what information is available from the government and then provide a way for them to link to that information directly.

The refocused GILS can be summarized as:

An easy-to-use and coherent government-wide information search service available from one or more service points that enables users to discover, locate, select, and access publicly available government information resources (e.g., agency information systems, specific information dissemination products, and existing locators to those products) through standardized metadata that describe those resources and provide direct links to the described resource (e.g., full-text documents, other online services).

The investigators view this refocused GILS not as a radical break with the current GILS initiative but rather as an evolutionary refinement to the concept of GILS.

4.3.9. RECOMMENDATION: Focus Scope of Descriptions on Network-Accessible Information Resources

Discovery and identification are logical prior steps to accessing or acquiring government information. Assuming that agency information resources are described by GILS in a manner that they can be discovered, the next challenge is for users to access or acquire the information described. This problem is compounded by the environment in which GILS is implemented.

GILS is a networked-based service. Since early 1994 when Web browsers became easily and freely available, Internet users have become conditioned to browsing and retrieving the full-text of electronic documents and being linked to online databases and other information services. The importance of this

"conditioning" cannot be underestimated when refocusing the GILS effort to assist discovering, identifying, and accessing government information. User input to the evaluation suggested strongly that simply providing a "virtual card catalog" of government information is not acceptable. A networked locator to resources is of far greater utility when the resources described are immediately available for access (e.g., one or two "mouse clicks" away).

Currently, GILS records describe both electronic and non-electronic resources. It is highly unlikely that non-electronic resources will be retrospectively digitized and made available online unless agencies see a benefit to doing so (e.g., a report that is in high demand, as a way to reduce the manual handling of documents frequently requested, etc.). One question that must be addressed in a refocused GILS effort is: what should be the scope and coverage of GILS?

One aspect of the coverage of GILS records is the extent to which GILS records will exist for *all* agency information resources. OMB Bulletin 95-01 directs agencies to create GILS records for three types of resources:

- Privacy Act Systems
- Automated information systems (AIS)
- "Locators that together cover all of [agency] information dissemination products."

An agreement between NARA, OMB, and GPO dealt with Privacy Act Systems (see Appendix A-4). A review of GILS records shows that agencies are describing AIS, but this study did not attempt to examine whether implementing agencies had created GILS records for all AIS (the purpose of this evaluation was not to address "compliance" in the audit sense of the word). The GILS record content analysis (see Appendix E-2) addresses the difficulty of understanding—from the description provided by GILS records—what "discrete set of information resources organized using information technology" (from definition of AIS in OMB Bulletin 95-01) comprise a particular AIS.

A more problematic area for producing GILS records is to list the "locators that together cover all of its information dissemination products" where "locator" is defined in OMB Bulletin 95-01 as an "information resource which identifies other information resources, describes the information available in those resources, and provides assistance in how to obtain the information." OMB Bulletin 95-01 uses the definition from OMB A-130 for information dissemination product as "any book, paper, map, machine-readable material, audiovisual production, or other documentary material, regardless of physical form or characteristic, disseminated by an agency to the public."

The review of GILS records done in this study indicates that some agencies are describing individual information dissemination products (e.g., a discrete publication or database), not simply "locators" that contain listings of those products. For some users, an "item level" description or granularity of the GILS records is much more helpful, especially when the item is in digital form and one can link from the GILS record to the actual item directly. More fundamentally, agency practice of creating GILS records that describe individual items reflects little understanding by agencies of what constituted "GILS Core" records, or possibly reflects the ambiguity of that concept. In addition, such practice may have been a response to the lack of agency locators that policy assumed existed and which were to be described by GILS Core records.

GILS assumed the existence of agency information locators, but, in fact, many agencies did not have a set of locators that cover "all of its information dissemination products." Agencies were then faced with the question: if no agency locators exist that cover *all* their information dissemination products, how should they proceed with their GILS implementation? Were they first to create the locators before creating GILS records that describe them? Or, could they simply begin using GILS to describe individual information dissemination products, whereby the GILS record itself became the "locator?" The creation of GILS records (i.e., the capture of metadata) at the item level for all the existing information dissemination products, however, would be resource intensive.

OMB 95-01 directed agencies to compile inventories if they did not exist. "As a first step, agencies should inventory their existing holdings and institute adequate information management practices.... By December 31, 1995, compile an inventory of its 1) automated information systems, 2) Privacy Act systems of records, and 3) locators that together cover all of its information dissemination products. Each such automated information system, Privacy Act system of records, and locator of information dissemination products shall be described by a GILS Core locator record." The policy, however, lacked specificity regarding what and how those inventories should be made available. There was clearly a missing step between the compilation of the inventories and the production of a "locator" to the inventoried items.

The issue of coverage is a difficult one for policymakers and implementors in determining appropriate guidance. Can the scope of a refocused GILS realistically cover all government information resources, especially if agencies do not have existing locators to their information dissemination products? Without additional resources, study participants agreed it is unlikely that the vast holdings of agencies will ever be described at an item level by metadata records.

If a refocused GILS initiative centers on networked information discovery and retrieval, the value of describing resources (locators, databases, automated information systems) that are not in digital form or network accessible is questionable. Focusing the coverage on government resources that can be linked to electronically (i.e., either in digital form or electronically accessible) may be a positive response to the expectations of users conditioned by the Web.

The investigators recommend that the scope of the refocused GILS should be on primarily supporting the discovery, identification, and access to online and networked resources, and preferably resources available or cast in terms of the Web. This recommendation responds to the increasing number of American citizens who operate in the networked environment and who are likely to want immediate, networked access to information

described in a refocused GILS. Anything less will create frustrations and raise questions as to the utility of the service. This recommended scope should not constrain individual agencies from describing non-digital resources, but at a government-wide policy and implementation level, GILS would be so focused.

The investigators realize that users will be interested in government resources even if they are not available electronically, but recommend this narrowing of scope for the refocused GILS. An accompanying recommendation, however, is that agencies be required to create, when none exists, network-accessible locators that describe non-digital, non-electronic, and non-network accessible agency resources.

The investigators recommend that the following two parameters guide a refocused GILS service:

- **Purpose:** Discovery, identification, and access of government information (i.e., not records management, information management, or other functions) through structured metadata records
- **Scope:** Descriptions of electronic resources that are publicly accessible, so that users can move from the metadata record to the "actual" resource.

Following from this, the refocused GILS should:

- Promote record creation describing existing and new publicly-accessible automated information systems (AIS), with the provision that users can link directly to those AIS via the Web (i.e., implement an interface between the Web and publicly accessible AIS through scripting mechanism such as the Common Gateway Interface [CGI], Java, or other alternatives).
- Require agencies to produce network-accessible locators that are described by GILS records.
- Point users to the GPO compilation of Privacy Act Systems.
- Encourage description of discrete information products where appropriate (e.g., high-value publicly accessible

documents such as the President's budget) and which are not covered by network-accessible agency locators.

This latter recommendation is problematic because of the resources it will take to create such records. There are several options, however, that can provide agencies with some flexibility:

- If there are machine-readable metadata records of agency resources held on internal, non-networked databases and servers, use an automated procedure to convert those records to standardized or compliant GILS records.
- Identify existing electronic locators to agency information resources and describe those in GILS records with a link from the GILS record to the locator.
- Identify frequently requested information dissemination products and describe those in GILS records, and ensure that those products are in digital form for network access and available via linking from the record.
- Identify all other information dissemination products that are in digital form (including resources available via an agency's web site) and describe those in GILS records, with links between the record and the information product.

A comprehensive list of government information resources is desirable, but if locators for all agency resources do not exist—especially given the current “do more with less” policy environment—a certain realism must be reflected in the refocused GILS policy.

Finally, and most importantly, agencies should be directed to create a GILS record for *each and every new* information dissemination product or ensure that such products are covered by agency locators in a timely manner. Determining the appropriate set of GILS record data elements needed for such item level description to support networked information discovery and retrieval is a question that needs to be addressed (see Section 4.4 for additional discussion of metadata). In

addition, government-wide and agency-level policymakers need to identify classes or categories of information dissemination products that deserve item-level description in GILS and develop government-wide guidance for agency implementors. Retrospective cataloging of existing resources may never be carried out in a comprehensive manner. Therefore, the investigators recommend a “from this date forward” policy that would require GILS records for new information dissemination products. This approach will, over time, populate GILS databases with records that reflect increasing coverage of agency products and resources. Further, since these resources and products begin life as an electronic file, an ever-increasing number of GILS records will be linked to digital copies of the products.

4.3.10. RECOMMENDATION: Identify Responsibilities and Authority for Policy Leadership, Government-Wide Coordination, and Oversight

For a refocused GILS effort to emerge and flourish as a *government-wide* initiative, the decentralized, distributed nature of the current approach needs to be balanced by some level of centralized oversight and coordination. Government-wide leadership of the refocused GILS initiative will be necessary. If OMB is unable to provide the leadership, coordination, and oversight, it must designate an appropriate body with such responsibility, and attendant authority and accountability. The goal is to establish formal mechanisms for addressing the refocused GILS initiative outlined in this report. The investigators view the GILS Board and the Chief Information Officers (CIO) Council as appropriate bodies to lead the refocused GILS effort.

The investigators recommend that the GILS Board—as an established body—has an important role regarding overall policy development and leadership for the refocused GILS effort. OMB Bulletin 95-01 provides a mandate for the existence of the Board, and the Board could be charged with responsibilities in addition to its current charge related to annual evaluation and reports on the progress of GILS. Current language in OMB Bulletin 95-01, “The Board may ask the heads of other agencies to designate representatives to serve on the Board or on

task forces established by the Board," enables the Board to create task forces that could assist in the refocused GILS initiative. The GILS Board should have the responsibility, authority, and accountability for formulating the policy direction for next phase of GILS development. OMB may be required to issue policy, but OMB should draw upon the Board's recommendation for the content of that policy. The investigators further recommend the following:

- GILS Board membership include representatives from the Small Agency Council and the CIO Council
- GILS Board establish a GILS task force consisting of representatives from Federal agencies as well as public users to refine and articulate the scope, purpose and goals for a refocused GILS.

The recently established CIO Council also has an important role to play in the coordination of GILS activities across the government. As an interagency body, the CIO Council could create one or more technical committees and working groups for discussions related to technical issues and concerns regarding GILS development. The CIO Council could, for example, establish a GILS Committee that would be responsible for government-wide coordination of the refocused GILS effort. Its working groups could address specific issues such as metadata record elements, marketing, ongoing evaluation, etc. The focus of CIO Council activities should be on technical and implementation concerns (as opposed to government-wide policy that the GILS Board would provide).

A CIO Council GILS Committee would provide a forum for Federal implementors of GILS to discuss and agree upon their requirements for the GILS Profile, which can then be taken to the GILS SIG for action. The CIO Council would be an appropriate unit for agencies to report their GILS implementation progress, and with such information the CIO Council could maintain the registry of known GILS implementation. Given its interagency makeup, the CIO Council would be an ideal forum for the identification and dissemination

of GILS "best practices" related to all aspects of GILS implementation.

Identifying a formal body as a home for technical and operational coordination responsibility, authority, and accountability should also provide increased credibility for the refocused GILS effort.

4.3.11. RECOMMENDATION: Implement a Refocused GILS Initiative

With the passing of the December 1996 OMB 95-01 deadline for GILS implementation and the conclusion of this evaluation study, GILS may be said to have completed its first phase. Pursuing a refocused GILS can be considered a second phase for the initiative. The question that faces policymakers—at both agency and government-wide levels—is how to take the next steps in evolving and implementing a refocused GILS that has the clear purpose of supporting the discovery, identification, and access of government information.

The Federal GILS initiative was driven in part by the Clinton Administration's efforts at reinventing government and the development of a National Information Infrastructure (NII). GILS, and its use of information technology, had the potential for supporting the accomplishment of agency mission by providing a mechanism for better information management (e.g., inventorying agency resources). Further, GILS was to support enhanced public discovery, identification, and access to government information. In Spring 1994 as the final GILS Profile specifications were being completed and the Information Infrastructure Task Force (1994) report on GILS was released, there were pressures to implement GILS as soon as possible. OMB Bulletin 95-01 directed agencies to begin developing their implementations in 1995.

In retrospect, the implementation would have profited from a GILS pilot program. Many of the issues encountered through this study could have been identified earlier, and with less onerous consequences, had a pilot program experimented with the various GILS requirements. As a case in point, the Canadian government established a GILS pilot project in 1996 and recently completed an evaluation

of the pilot (see Appendix I for copy of the report on the Canadian pilot).

A U.S. Federal GILS pilot program would likely have identified the following issues:

- **Record Creation:** How much effort would it take to compile the information needed to create records? What barriers might be encountered? What data input mechanisms could be devised to ease the burden of data collection and data input?
- **Z39.50 Software:** What was available and what would be the demands for implementing Z39.50? Were the GILS Profile specifications realistic and implementable?
- **Record Content:** Had appropriate data elements been defined? Were the data elements and the content of those elements clear and usable?
- **Records Management:** How would GILS records support records management? To what degree did GILS metadata elements satisfy records managers information requirements?
- **Usability of GILS:** What was the best way to present GILS data to users? To what extent did it satisfy users?

A pilot program could have not only identified problems and issues, but could have served as a testbed to resolve them.

Many agencies are not only skeptical about GILS after the past two years; some are frustrated and angry from trying to do GILS with no new resources and little realization of tangible benefits from their activities. Exhortations from policymakers will not be enough to overcome resistance to doing anything more with GILS as it currently exists (either at a management or staff level). A refocused GILS must demonstrate that it can solve networked information discovery and retrieval challenges and provide real benefits to agencies and their users.

Assuming that redefinition of GILS occurs along the lines recommended by the investigators, that

the purposes and objectives for a refocused GILS are identified and articulated, and that organizational units are delegated with the responsibility, authority, and accountability for coordinating a refocused GILS initiative, the next step should be the implementation of a phase two GILS pilot program. A GILS pilot program could be used to implement the recommendations offered in this report.

For the refocused GILS, a period of time (e.g., 9–12 months) should be allotted to a pilot program. During this period, a small selected group of agencies could participate in pilot implementations of GILS that address some of the specific issues and problems identified in this evaluation. Agencies should be chosen that reflect differing missions, sizes, information holdings, levels of information management sophistication, etc. Reasons for conducting a pilot program include:

- Demonstrate that GILS improves public access to government information
- Demonstrate the tangible benefits to agencies
- Demonstrate the costs incurred by agencies
- Demonstrate an approach that improves user satisfaction in discovering and accessing government information
- Demonstrate the appropriate staffing required for successful implementation of GILS
- Demonstrate the technology solutions for record creation, information retrieval, record presentation, etc.
- Demonstrate how GILS can be integrated into other agency information handling processes
- Document how GILS can be implemented and share lessons learned, best practices, etc.
- Showcase the potential of GILS in improving information discovery and access both for agencies and users.

This pilot program assumes that policy leaders, project-management and technical experts, and input from various user communities have refocused goals for GILS, have identified specific and measurable objectives for GILS, and have provided guidelines for implementors to follow. The success of the GILS

pilot program can then be gauged against (1) conformance to specified goals, objectives, and requirements and (2) user feedback as to the degree to which GILS "enables" information discovery, identification, and access.

The refocused GILS policy should communicate clearly the goals or future conditions so that agencies and users can envision the purpose, scope, and utility of GILS. Agencies should support these goals by developing specific, realistic, and time-phased objectives with assigned responsibilities, accountabilities, and authorities (see Appendix G for characteristics of successful objectives). This approach can encourage measurable performance, and the goals and objectives—and procedures for measuring and assessing performance—would provide a basis for agencies to comply with *The Government Performance and Results Act of 1993* (GPRA) requirements related to GILS activities.

4.3.12. RECOMMENDATION: Require Agency Reporting on GILS Progress and Reward Agencies That Achieve Stated Objectives

Existing U.S. Federal GILS policy lacks a requirement for agencies to report on the progress of their GILS implementations. The GILS Board is charged with conducting yearly assessments on the progress of GILS and documenting its findings in an annual report. Without any agency reporting requirements, how the GILS Board would gather information for its annual assessment is unclear. In general, neither sticks nor carrots are identified to "encourage" or "reward" agencies for their progress (or lack thereof).

The evaluation study also identified a lack of incentives and benefits to agencies that participated GILS. The incidence of disincentives may be higher than that of incentives. Agencies, especially smaller agencies, view GILS as providing little return on investment (i.e., much burden, few benefits). In some cases, especially where records managers are charged with GILS record creation, there are disincentives for creating GILS records. Not only did the records

managers have to create records, but, if those records described unscheduled items, the items then had to be scheduled as well.

Agencies that demonstrated creativity and innovation in their GILS initiatives received no public recognition. Nor were there any rewards for or acknowledgment of agencies that met the deadlines of OMB Bulletin 95-01.

If the GILS Board and the CIO Council assume specific responsibilities (outlined above) for the refocused GILS initiative, they will need adequate information from agency implementors to manage and coordinate the initiative successfully. The investigators recommend that agencies be required to report at least annually on the status of their GILS implementation. Such reports should include the following:

- Network address of the agency's GILS records
- Implementation used for providing network access to the records including type of database and search engine used
- Number of GILS records created in reporting period
- Total number of GILS records created
- Number of GILS site accesses, searches, and record retrievals per agency log analysis
- An estimate of the percent of agency information resources described by GILS records per scope of the refocused GILS initiative
- Identification of any evaluation/assessment conducted by the agency of its GILS implementation
- Identification of mechanisms employed to gain user input into development of the agency GILS.

The investigators also recommend that policymakers (e.g., OMB, the GILS Board, and the CIO Council) explore the creation of incentives for agency compliance and develop a program of rewards or public recognition for those agencies that demonstrate creativity in accomplishing and/or exceeding the clearly stated objectives of the refocused GILS initiative.

The proposed GILS pilot program needs to be authorized or supported with a source of money. The *Information Technology Management Reform Act of 1996* (ITMRA) established an Information Technology Fund (including a proposed funding source) consisting of an Innovation Loan Account Fund (to be funded out of existing agency IT budgets) and a Common Use Account Fund (to support multi-agency acquisitions). Some of these funds might be tapped, on a reimbursable basis as required with the Fund, to support innovative development of GILS efforts. This could be done on a proposal basis, whereby agencies could submit short proposals for innovative projects that address significant problems with GILS and solutions for which can have government-wide application, and could be awarded funds to carry out innovative projects. Challenges and problems identified in this report that would be suitable for such pilot program activities include the efficient capture of metadata in conjunction with electronic document management systems, usability studies for options in presentation of GILS records, identification of high-value metadata elements that support discovery and retrieval of government information resources. Government-wide solutions for improving public access and agency information management resulting from the use of the IT Fund appears clearly justifiable.

4.3.13. RECOMMENDATION: Ensure Ongoing, User-Based Evaluation for Continuous Improvement

The investigators spoke with many agency staff who are committed to GILS and who are making good faith efforts in implementing it even if they do not have adequate resources allocated to their work. Yet, with notable exceptions where agencies (e.g., EPA) actively solicited potential users' input, users external to the agencies have not been involved in GILS design and implementation. GILS has been dominated by agency, resource, and system-centered considerations. The online user assessments of GILS highlighted that, overall, it is not a user-centered system.

One key finding from this study is that a number of evaluation and self-assessment tools can be used by Federal agencies to assess the overall success of their GILS efforts. A by-product of the study is the development and testing of techniques and instruments that are reprinted in the appendices. Techniques such as server log analysis, user scripted assessment of a GILS site, record content analysis, as well as more familiar focus groups, surveys, and interviews provide important indications of the overall health of GILS.

While the investigators heard agency representatives lament the lack of time and resources for assessment, especially user-based assessment, an ongoing evaluation of GILS is essential if it is to improve networked information discovery and retrieval of government information. The study finds that mechanisms will be needed to conduct both government-wide and agency-level assessments in the next phase of GILS. A number of those mechanisms and data collection instruments should be adapted from this study.

The investigators recommend that agencies establish ways of routinely seeking user input on the design and implementation—as well as the criteria for determining success—of the refocused GILS (and other public access activities). The CIO Council, as a newly constituted coordinating and policy body for GILS, can lead this aspect of the GILS initiative by identifying procedures and practices to solicit and capture a wide range of user perspectives.

User involvement should begin during the phase of clarifying the purpose, goal, and objectives for the refocused GILS (e.g., having public representatives on the GILS Board and its GILS Task Force). The proposed GILS pilot program must build in user involvement (e.g., early input into the design and specification), and user-based evaluation should be ongoing through the pilot program activities and implementations. For example, public and government documents librarians could serve as important sets of users in assessing and evaluating GILS clients that could be developed and tested as part of the pilot program.

4.4. OPPORTUNITY: IMPROVE GILS EFFICACY IN NETWORKED INFORMATION DISCOVERY AND RETRIEVAL (NIDR)

GILS is a networked-based service that can assist users in discovering and accessing government information. In the early 1990s, the term networked information discovery and retrieval (NIDR) emerged to describe the complex activities and problems—technical, organizational, and users—involved in search and retrieval in the Internet environment. GILS serves as an example of a NIDR system. Findings from the study indicate that GILS utility as a mechanism for users to discover, locate, select, and access government information is limited. Table 4-5 summarizes a series of findings and recommendations related to this aspect of GILS.

Section 4.3. discussed issues that require policy, administrative, and organizational attention. The issues related to NIDR are, however, of a different order. In many respects, NIDR is a research area in which computer and information scientists are framing and addressing difficult challenges related to distributed search and retrieval, the character and utility of metadata, interface design, and others (see Lynch, et al., 1995). The many digital library projects underway provide environments where many of the issues and challenges are becoming more clearly defined. Scalable solutions to some of the problems have yet to become operational. The findings reported here from the implementation experience with U.S. Federal GILS will contribute to the understanding of some of the NIDR problems. Given this situation, some of the recommendations should be viewed as the investigators' indication of potential next steps. Further, the findings and recommendations point to additional research that needs to be carried out, and Chapter 5 identify the major research topics related to GILS and NIDR.

Table 4-5
Improve GILS Efficacy in Networked Information Discovery and Retrieval

OPPORTUNITY: IMPROVE GILS EFFICACY IN NETWORKED INFORMATION DISCOVERY AND RETRIEVAL (NIDR)	
Findings	Sources of Evidence*
4.4.1. Web Technology Has Raised Questions about the Role of GILS	FG, SU, SV, US
4.4.2. GILS is an Agency-Centric, Rather than Government-Wide, Service	FG, SV, US
4.4.3. GILS Metadata Are Difficult to Capture	CA, FG, SV
4.4.4. Limited Updating and Maintenance of GILS Records	CA, FG, SV,
4.4.5. No Clear Agreement on Adequacy of GILS Record Data Elements	CA, FG, SV, US
4.4.6. Different Types of Resources Represented in GILS Records	CA, FG, SU, SV, US
4.4.7. User Reaction to GILS Is Not Positive	FG, SU, SV, US
4.4.8. GILS Record Display Varies Widely and Is Criticized by Users	CA, FG, SV, US
4.4.9. User Orientation and Instruction is Inadequate	FG, SU, US
Recommendations	
4.4.10. Continuously Evaluate GILS Policies and Standards against Emerging Technologies, Especially the Web	
4.4.11. Specify Resource Types And Aggregation Levels	
4.4.12. Enforce Consistent Use Of Metadata That Are Empirically Demonstrated to Enhance Networked Information Discovery and Retrieval	
4.4.13. Improve Presentation of Metadata	
4.4.14. Develop Policy and Procedures for Record Maintenance	
4.4.15. Promote Interagency Cooperation and Use of GILS for One-Stop Shopping Functionality	

* CA=content analysis of GILS records; FG=focus group sessions; KP=interviews with key participants;
LA=log analyses of Web servers; SU=survey conducted at the 1996 GILS Conference;
SV=site visits to selected agencies; US=scripted online user assessments of GILS

4.4.1. FINDING: Web Technology Has Raised Questions about the Role of GILS

Web technology has developed rapidly during recent years. The degree to which Federal agencies embraced the Web as a means for providing access to government information resources, disseminating information products, and providing a range of information services during that time could not have been foreseen at the time of GILS development and the writing of OMB Bulletin 95-01. DiCaterino and Pardo (1996) provide an analysis of the ability of Web technology to provide a universal interface to government information. At issue is how GILS can best take advantage of the Web technology while providing an essential service not currently offered by Web technology—namely, a search and discovery service.

Currently, all known U.S. Federal GILS implementations are accessible via a Web interface. Yet the study found a certain amount of confusion, if not contention, between the roles of GILS and uses of agency Web site. Agency officials also had varying opinions as to what GILS records are supposed to describe versus what Web pages should include and describe. To a large degree, agencies are still experimenting with how best to integrate these two approaches for information access and dissemination. Interestingly, 79% of respondents to the GILS Conference survey agreed with the statement: *Every agency Web homepage should have a link to the agency's GILS.* Only 16% of respondents agreed that: *The World Wide Web reduces the need for GILS.*

In part, the confusion stems from a lack of understanding of two key elements of GILS:

- Structured metadata (i.e., GILS records) that describe agency information resources
- Z39.50, the information retrieval protocol.

The GILS records, as structured metadata, provide a standard way to describe agency information resources in a semantically consistent way (see Appendix F). More importantly, Z39.50 provides for "semantic interoperability" in that it enables client

software to precisely express a search query to multiple search engines and supports the retrieval of complex records (Lynch, 1997; see also Lynch, 1992). Z39.50 servers and clients that support the GILS Profile share an understanding for search and retrieval, and according to Lynch (1997), "Z39.50 provides maximum leverage [for search and retrieval] where there is a shared understanding between client and server of rich and specific information semantics." Thus, the GILS records and Z39.50 provide an important basis for searching across multiple databases and servers.

A number of study participants suggested that Web search engines provide sufficient searching power. Yet, Web search engines are limited, based as they are on a simple model of retrieving HTML documents from multiple sites and building large centralized indexes based on the occurrences of words in the HTML documents. The search engines are very powerful and robust for full-text searching of HTML documents. However, users cannot search, for example, for a copy of the document with the title of "Circular A-130, Management of Federal Information Resources" and published by the Office of Management and Budget, and be assured that the results that come back from the search engine do not contain commentary on A-130, email messages about A-130, and bibliographic citations to A-130. Another drawback to the Web search engines is that they do not "see" all electronic resources that may be network accessible. For background on Web search engines and their capabilities, see Koster (1997) and Liu (1996).

A more critical area where Web search engines provide only limited service in discovering and identifying resources is electronic databases. Since Federal agencies' databases are important and valuable resources, GILS provides an important function by enabling standardized descriptions of these resources that are only slightly "visible" to the Web search engines. In many cases, what is visible to the Web search engine is not the database itself but usually an HTML page (possibly forms-enabled) that the user interacts with to use the database.

The following example helps illustrate this point. The Security and Exchange Commission (SEC) makes available its Electronic Data Gathering,

Analysis, and Retrieval (EGAR) database through a Web site www.sec.gov/edgarhp.htm. A web search engine could index the EDGAR web site but would limit its indexing to words appearing on the site. The SEC has created a GILS record for the EDGAR database (available on GPO's GILS service), and the GILS record provides structured information including time period of content of the database, its purpose, how to request information, and other useful information. Using a well-known Web search engine (AltaVista) and the search terms EDGAR and "Securities and Exchange Commission," the search engine found "8,000 documents matching the query." However, none of the first 20 "hits" pointed to the EDGAR homepage. Further, even if the homepage would have showed up in the result set, the listing would not have provided the type and scope of information contained in a GILS record. One particular hit pointed to "EDGAR Online" <www.edgar-online.com/>, a commercial service provider of SEC information; in fine print at the bottom of the page for EDGAR Online, there is the following disclaimer: "EDGAR ONLINE is a product of Cybernet Data Systems, Inc. and is neither approved by, nor affiliated with the SEC." Compare the lack of results when searching a Web search engine with the results when submitting the same search query on GPO's GILS. In the latter case, the GILS record for the official SEC EDGAR database was near the top of the result set list, plus it provided additional authoritative information from the originating agency.

Hamner and Favaro (1996) identify a potential synergy between the Web and Z39.50 by acknowledging their separate strengths. The Web provides hyperlinks between systems and documents types, as well as a relatively easy mechanism for publishing and an interface to existing databases. The strength of Z39.50 is structured searching and document discovery, precisely the goal of GILS.

The challenge for the future is to refocus GILS efforts to emphasize the discovery of government information provided through the GILS records and the structured searching provided by Z39.50. Once users discover the information resources by searching GILS, it is necessary to provide seamless links from GILS metadata records to individual documents (in

full-text) or other electronic resources accessible that may be available on agency Web sites.

One critical result of the Web's influence on GILS is the increase in users' expectations of being able to access the full-text of documents and other electronic resources. Not satisfied simply with viewing "pointer" or descriptive records, users want access to the "actual" information resource. Users participating in the online assessment of GILS expressed "disappointment," "surprise," and "confusion" to the absence of full-text (i.e., the actual documents) when interacting with GILS implementations.

The investigators maintain that producing quality metadata is an important contribution of GILS. Metadata, however, may not be sufficient to satisfy users' information needs. The GILS record structure provides data elements to enable linkages to the information resource described in GILS. Many agencies are making an effort to use this feature to take the user from the record to the resource described (e.g., *The Budget of the United States Government* (OMB) and GPO's *Monthly Catalog*). In the record content analysis component of the evaluation, approximately 25% of GILS records examined featured at least one instance of hypertext linkage. While linkages occurred most frequently in the Available Linkage data element (approximately 15%) and thus enabled linkages to the resource described, instances of hotlinks were also present in fields such as the Distributor Network Address and Abstract data elements (as well as some locally defined elements). While the maintenance burden of hypertext is recognized, users' expectations for it will continue to accelerate for the foreseeable future.

Some agencies have integrated GILS into their Web site by providing a link to the agency's GILS on the agency homepage. At EPA, GILS records assist Web visitors navigate the Web site to find information, even though the records are not labeled as being part of "EPA's GILS." Most study participants thought that GILS should have a more discernible relationship with an agency's web site. They wanted to integrate GILS better with agency home pages and with other information systems and information product catalogs. How specifically this should be done is a matter of some considerable debate.

The study finds that agency Web implementations have not replaced the need for GILS or a GILS-like service. Available Web search engines that index Federal Web sites and search engines on individual agency Web sites do not provide access to *government-wide* finding tools, catalogs, or indexes across agencies and across related topics. Nor does the Web supplant the power that Z39.50 offers for interoperable search and retrieval. Finally, many resources of the government, such as electronic databases, are not "visible" to Web search engines. Even if a search engine indexes an interactive form page for a database, the value-added, structured information captured in a GILS record is not available for the user.

4.4.2. FINDING: GILS is an Agency-Centric, Rather than Government-Wide, Service

The study finds that the Federal GILS initiative has not resulted in a *Government-wide* Information Locator Service; rather, it has resulted in separate *Agency* Information Locator Services (AILS). Agency GILS that have been implemented are confined almost exclusively to resources within a particular agency. Until recently, users could conduct cross-agency search and retrieval capability when searching the GPO and FedWorld GILS sites, but the searches were limited to agencies which had contracted with GPO or FedWorld to mount their GILS records. In April 1997, GPO announced it had implemented an application where a user could submit a query across agencies' GILS records whether or not GPO had mounted those databases of GILS records on its site.

In the online user assessment sessions conducted as part of the evaluation, users were nearly unanimous in their agreement that *all agencies' GILS should be searchable together, from one Web site as well as all government documents on the Internet should be hotlinked from one electronic card catalog* (see summary of user sessions in Appendix E-3, specifically Question S32a).

OMB Bulletin 95-01 recognized the need for the U.S. Federal GILS to be built from agency components. The vision of GILS reflected a

decentralized collection of agency information locators. It specified two approaches, however, for creating a logically centralized albeit physically decentralized government-wide locator. First, GILS servers were to implement the Z39.50 protocol, which would allow a single Z39.50 client to interoperate with all GILS servers and provide an impression of seamless searching and navigation among those distributed servers (see Lynch). While fully compliant Z39.50 GILS servers are being implemented, the incidence of desktop GILS clients has been relatively low.

Most users connect to GILS servers through a Web interface (e.g., a Web/Z39.50 gateway), which limits users to searching GILS records that are offered through the gateway. Without Z39.50 GILS clients that provide users the capability to search across one or more GILS servers, cross-agency searching has yet to be achieved. (Nor does the user have the control over the display and views of GILS records that Z39.50 affords.) "Integrated" services, such as those offered by GPO and FedWorld, are important, however, as they provide users with some modicum of government-wide searching.

Second, GILS record creators and maintainers could include cross references to other resources that might be of interest to a user, whether from the originating agency or resources at other agencies. The identification of these related resources would allow a user to link to or search for these resources that were themselves described by other GILS records. In the record content analysis carried out as part of the evaluation (see Appendix E-2 for the analysis), the occurrence of cross references in GILS records was negligible. Given the difficulty for many GILS record creators to gather agency information to create GILS records, it was probably unrealistic to assume that agency staff would go the additional step in referencing related resources, especially those of other agencies. However, this capability, along with realistic procedures for maintenance of cross-references, should be a goal of the refocused GILS initiative. One step in this direction would be the development of criteria to help identify suitable resources that could be cross-referenced.

Other models of cross-agency searching or access to agency resources are available. First is the brokered-

GILS model with a single agency (e.g., GPO and FedWorld) providing a single point of access for searching against more than one agency's GILS record simultaneously. As noted earlier, GPO has mounted 27 agencies' records and allows a user to search across all records with one query. FedWorld provides searching of three agencies' records at one time. While this model moves in the direction of a government-wide service, it is based on a model where agency databases of GILS records are hosted at a centralized site. Searching is limited to the agency records available at that centralized site.

Another model is represented by the Advanced Search Facility (ASF) effort. This interagency initiative has been developed under the leadership of the Department of Commerce and is informed by recent Web models of search and retrieval. Web search engines provide for the centralized and automatic indexing of resources accessible by Web robots (Finin, 1997). The robots "crawl" the Web, pull documents to the indexer, index the documents, and then offer a search service against the centralized indexes. Examples of such Web search engines include AltaVista, Yahoo, and Excite. The user connects to a search engine, submits a search which is then executed on the indexes, and is given a list of resources that "match" the search criteria. The user then links to the resources of interest. This model is also based on centralization of resources—in this case, the centralized indexes built by the web robots and search engines.

The objectives of the ASF initiative address the problems of searching for information across many agencies. The ASF expands the indexing of networked resources beyond the Web resources currently covered by the Web search engines, and distributes indexing responsibilities to the distributed servers. Discussions with staff working on the ASF indicated to the investigators that the ASF appears to have potential to help solve the GILS problem of cross-agency searching. At the time this report is being written, a Request for Proposal (RFP) for the ASF has not been issued. Thus, specifics of the project and how it might assist GILS may be premature.

A third model is represented by interagency initiatives that use the Web to provide access to topical or subject-oriented collections of government information and services. Examples include:

- Business Advisor, the one-stop electronic link to government for business
<http://www.business.gov/>
- Federal Statistics Initiative
<http://www.fedstats.gov>
- National Environmental Data Index (NEDI) <http://www.nedi.gov>.

Except for NEDI, the use of GILS to support such interagency efforts is not clear. These models do not provide a government-wide locator service as envisioned for GILS. And, as noted in the previous section, the Web does not provide a systematic solution to the information discovery problem. Topically-based resources must first be "discovered" by users before they can be used (i.e., describing these resources in GILS records would be appropriate). Further, such topically-based resources "pre-select" resources for users (which may be entirely appropriate). An analogy would be a special collection within a larger library. GILS provides a means to discover what is in the larger library.

Distributed information search and retrieval in the networked environment is a difficult problem—both technically and organizationally. Like other aspects of the GILS initiative, cross-agency searching using Z39.50 appeared reasonable; to date, effective government-wide searching for government information has not been achieved, either through GILS or any other mechanism. The agency components of GILS, however, are a vital foundation for a government-wide locator. Without mechanisms such as the centralized point of access via a centralized service that actually mounts all agencies' GILS records (e.g., GPO and FedWorld) or centralized index (e.g., a Web-like search engine or ASF) or the deployment of compliant Z39.50 GILS servers and clients, GILS will remain a distributed, unconnected set of AILS.

4.4.3. FINDING: GILS Metadata Are Difficult to Capture

At the core of GILS is the standardized record with defined data elements that can be used to describe agency information resources. Although a number of study participants indicated a limited understanding of the concept of "metadata," others view the standardized record as offering GILS' most valuable contribution for enhancing discovery and access to government information. The investigators remain convinced that a standardized set of metadata elements is one of the clear strengths of the GILS initiative.

The term "metadata" has evolved into common usage in the networked environment to describe "data about data." That definition is accurate, but its helpfulness is limited. The investigators determined that while people used the term "metadata" often in regard to GILS, common understanding or agreement on what specifically was meant by "metadata" when discussing GILS was not readily apparent. Many study participants were not clear about or had an appreciation for the role of metadata in networked information discovery and retrieval. A number of study participants suggested that Web search engines replaced the need for GILS and GILS records. Such a view is incorrect and suggests a need for better training about the use and benefits of metadata in NIDR (see Appendix F for brief discussion on the role of GILS metadata in NIDR).

Study participants were concerned with the cost-effective capture of data needed to create GILS records. Although some agencies, such as DoD, have implemented an online process for creating GILS records, a prior step of gathering the information to put into the record is necessary. Agency staff involved with record creation pointed out the difficulty in gathering that information. While the Office of Primary Interest (i.e., the staff or office responsible for a particular agency resource) may have the pertinent information about a resource to be described in a record, agency staff indicated that cooperation from those offices was not always enthusiastic. The effort in gathering GILS record information should not be underestimated. As

currently done in most agencies, GILS record creation is time consuming and requires major effort.

One wonders why more agencies did not make use of freely available record creation and data input aids such as DTIC's electronic input form. Part of the answer may lie in uncertain or unfamiliar lines of communication among agencies (e.g., civilian and military), and part of the answer may be that the aids were not known to be available at the time when agencies had to make decisions regarding input procedures. The technology infrastructure or local expertise within an agency were also constraining forces in using such software applications. Better cross-agency coordination could have led, however, to substantial government-wide efficiencies in records creation.

Some agencies preferred to centralize data entry. These agencies believed that they achieved greater record quality assurance in this fashion. On the other hand, centralization sometimes complicated the process of updating and maintaining records, since the people closest to the information resources would need to go through the central point for record updates. The practice of centralization becomes problematic when the described resource or its descriptive metadata change frequently; however, implementation of an updating schedule to allow periodic incorporation of changes may improve efficiency.

Overall, the study found that record creation at the time of the creation of an information resource is rarely done, that "best practices" for GILS records creation should be identified and publicized, and, overall, that the GILS record creation process should be simplified.

The study also found that agencies lack staff, funds, and other resources to retrospectively "catalog" their information. A number of study participants suggested that retrospective cataloging to create GILS records for "comprehensive coverage" an agency's information resources is unlikely. In part, this is due to the costs involved. Although this study did not attempt to collect information about costs in creating GILS records, it is reasonable to estimate such costs based on the costs involved in cataloging materials in libraries. Recent data from the Library

of Congress suggest that cataloging a single monograph can range from \$25 to over \$100, depending on the depth and extent of cataloging and classification.

Study participants offered suggestions for some form of automatic capture of metadata at the time of creation of the information resource. New tools and procedures such as electronic document management systems (EDMS) provide one possible scenario for capturing metadata, at least of document-like objects.

4.4.4. FINDING: Limited Updating and Maintenance of GILS Records

One important question concerning the quality and timeliness of GILS records relates to the maintenance activities of updating, verifying, and ensuring record accuracy and currency. Without ongoing maintenance, the quality (e.g., accuracy and currency) of GILS records will degrade. Inaccurate, out-of-date records will not improve access.

Updating and maintenance burdens will vary based on a number of factors. One factor is whether the records are mounted locally on an agency server or mounted on a host agency server. For example, GPO staff mount the records "as submitted" and rely on each agency to notify them of record changes, updates, deletions, etc. DTIC, on the other hand, has a procedure in place that "strongly encourages" agency maintenance of existing records on a regular basis.

Updating and maintenance may also be a function of the agency network infrastructure: do the GILS record creators have network access to check and correct the information contained in records they create? Finally, ongoing maintenance of GILS records will depend on agency staff perceptions of GILS' value. In a number of agencies, the lack of tangible benefits to date provides sufficient reason for them to say, "we created records, but we aren't going to put any more resources into the effort or maintain the records we created."

Some agency records officers responsible for GILS activities reported that once they created the original record they believed their job completed. Other

GILS records creators are dependent on others in the agency for updates and find little cooperation for obtaining the updated information. Still others told the investigators that they have neither plans nor intent to update the records created to date. Overall, the study found a lack of procedures and a general lack of interest at many agencies in updating and maintaining GILS records.

4.4.5. FINDING: No Clear Agreement on Adequacy of GILS Record Data Elements

Study participants noted that the GILS records may be the lasting contribution of the U.S. Federal GILS initiative. By this, they meant that across government agencies, staff used a standardized set of data elements to describe agency information resources. This effort is analogous to the evolution of a standard bibliographic entry in library catalogs. Standardized, structured metadata records such as GILS can have a longevity beyond the life cycle of the access systems on which they were initially implemented. The GILS records can be viewed as platform and application independent, and the investment made in creating GILS records can have long term payback. The structured records can be converted and migrated to other systems and other applications.

There was not agreement, however, on the adequacy of the GILS record, and in some cases, study participants questioned the usefulness of the many data elements defined for use in GILS records. Some record creators thought that GILS records required too much information, and they concluded that the cost of collecting the information outweighed the benefits of including that information in the records.

In some cases, the GILS records do not contain adequate data elements to support functions expected of GILS. Specifically, records managers participating in focus groups suggested that the information they need for record scheduling cannot easily be put into existing GILS data elements. They also thought that GILS was not an appropriate records management tool and were not interested in trying to "enhance" the data elements to the extent necessary for GILS records to be useful in records management.

The issue of appropriate metadata elements that support information discovery and retrieval warrants additional research, and such research is part of the larger research issues related to networked information discovery and retrieval (see Chapter 5).

Finally, a number of study participants were unclear as to who had or should have authority for the data elements for GILS records. Currently, the GILS SIG is responsible for maintaining the GILS Profile, in which the data elements are defined, and as such, the GILS SIG is open to any implementors of the GILS Profile and is not limited to Federal agency GILS implementors. No Federal forum for U.S. Federal GILS implementors exists where agency GILS implementors can discuss and review *their* needs regarding GILS data elements.

4.4.6. FINDING: Different Types of Resources Represented in GILS Records

Considerable discussion occurred in a number of site visits and focus group sessions regarding the types of resources and the granularity and/or aggregation of agency information resources represented by GILS records. The GILS record content analysis and the scripted online user assessment also identified issues regarding the unit of information described by a single GILS record (see Appendix E-2 for discussion of granularity and aggregation). Should agencies create GILS records for individual maps, publications, and documents? For individual databases, which may aggregate many discrete resources?

As stated earlier, OMB Bulletin 95-01 identified three types of information resources GILS records should describe:

1. Automated information systems
2. Privacy Act systems of records
3. Locators that together cover all of its information dissemination products.

Yet, based on discussions with agency staff and the GILS record content analysis, there is a great deal of uncertainty as to the appropriate level of granularity, or extent of aggregation, for GILS records. Some participants, for example, told the investigators that

they plan to produce GILS records for individual documents and resources because they were "key items" in their agency. Other agencies are creating records for collections of hitherto individual documents (e.g., aggregating "press releases" to be described by a single GILS record). Without government-wide guidance, agencies now have wide latitude for determining what resources and what level of granularity their GILS records describe. The result—from a *government-wide* perspective—for users is uneven levels of description and inconsistent representation of resources. This also results in users being uncertain as to the scope and coverage of a particular agency GILS based on the number of records that have been created.

The number of records created by an agency may or may not be an indicator of the degree of resource aggregation. For example, EPA has created approximately 240 GILS records; the Social Security Administration has created more than 1200. Is Social Security Administration creating too many records (they have many GILS records that describe one form) or is EPA creating too few? Absolute numbers of records are less helpful than understanding two important issues related to the GILS records:

- The granularity/aggregation of described resources (i.e., the extent to which individual information products are "collected" for description by a single GILS record)
- The overall coverage of information resources (i.e., the extent to which an agency's GILS records describe all agency AIS, Privacy Act systems, and locators per OMB Bulletin 95-01, or describe individual information dissemination product).

EPA had preexisting locators to much of its information resources, and by creating GILS records that describe those locators, EPA may be able to provide good coverage of its information resources through a relatively small number of records. If, on the other hand, an agency does not have existing locators and it chooses to describe individual documents and publications in individual GILS

records, then a larger number of records may be necessary to gain adequate coverage.

Another consideration is whether agencies use the same definition or criteria to determine what, specifically, constitutes an agency "resource" or "product" that should be described by a GILS record. The evaluation study's record content analysis developed criteria and procedures for assessing GILS records (see Appendix E-2), and, identified various types of resources described in a sample of approximately 80 records from all GILS sources. Table 4-6 summarizes the findings from this analysis.

Granularity and aggregation are not simple concepts. The record content analysis used the following operational definitions to deal with the issues of record aggregation:

- **Record aggregates object:** The GILS record, by virtue of its creation, collects discrete information resources that the record content indicates would not have otherwise been collected or aggregated (e.g., "General Files," "Press Releases," and "Forms").
- **Aggregated object represented:** The GILS record represents an *a priori* or purposeful collection of information resources (e.g., "Woodpecker

"Database" or an agency Web site). The GILS record represents an object that collects, or comprises, two or more discrete information objects, and that object represents a collection of standalone information files or products packaged together on the basis of a common theme or subject for functional convenience (e.g., a CD-ROM of regulations, a system of Privacy Act records, or a voice recording of employment opportunities).

- **Discrete object represented:** The GILS record describes a standalone document-level entity that does not meet the criteria for "object aggregates metadata" below (e.g., an Annual Report or a videotape).
- **Object aggregates metadata:** The GILS record describes a pre-existing metadata collection, or "locator," as an information resource (e.g., directory, catalog, or index).

Based on these operational definitions, Table 4-7 provides a summary of aggregation characteristics of information resources found in the sample analyzed. An important finding from this study is that agencies use GILS to describe collections of information resources not previously described. For example, a GILS record describing an agency's "press releases" (or some subsets of press releases) provides users with the opportunity to discover the existence of these resources.

Table 4-6
Resources Described by GILS Records

OBJECT REPRESENTED	N	%
Subject Matter Database	18	22%
Publication	16	19%
Miscellaneous Documents in <i>Ad Hoc</i> Collection	14	17%
Agency Homepage	8	10%
Organization	6	7%
Form	4	5%
Administrative Catalog	3	4%
Bibliographic Database	3	4%
Publications Catalog	4	5%
System of Systems	3	4%
Program	2	2%
Job Line	1	1%
Unknown	1	1%
TOTAL	83	100 %

Table 4-7
Aggregation of Resources Described by GILS

AGGREGATION	N	%
Record Aggregates Objects	30	36%
Aggregated Object Represented	21	25%
Discrete Object Represented	17	20%
Object Aggregates Metadata	10	12%
Unknown	5	6%
TOTAL	83	100%

Study participants could not define an optimal or appropriate level of granularity. Many concurred that existing GILS records describe a wide range of resources, of varying levels of aggregation, and that this phenomenon could affect GILS usability. Users indicated difficulty in knowing what to expect to find in GILS. Indeed, most in the online user assessment disagreed with the statement: *It is clear to me how agencies choose what to include in GILS.*

Currently, there are differing views of the level of granularity that is appropriate for inclusion in both the GILS records and for the items to be included in the GILS database. The result of these differing views is inconsistency in agencies' GILS records regarding the types of resources included and the detail of the descriptions for the resources. The study finds that specific guidelines are needed to clarify the types of information resources that should be described by a GILS record.

4.4.7. FINDING: User Reaction to GILS Is Not Positive

Throughout the evaluation study, the investigators heard little in the way of positive experiences from people attempting to use GILS for finding information. To capture user perceptions about and reactions to GILS concepts and serviceability, the evaluation featured an exploratory technique based on a set of scripted service encounters (see Appendices C-5, D-5, and E-3 for a description of the technique, the instrument (script), and results, respectively). In this simulation of how users might

use and assess a GILS, 10 undergraduate and graduate students at the University of North Texas and at Syracuse University completed a series of browse, search, and retrieval activities. Overall, users were confused and disappointed with the experience for a number of reasons, including:

- An inordinately high degree of user sophistication is required to exploit GILS (e.g., one user remarked "shouldn't have to feel like they're hacking into a government system" and another asked, "would you turn a twelfth grader loose on GILS?").
- Users were interested in and/or expecting to gain access to full-text.
- GILS records were hard to read, contained unnecessary information, and were not linked to the actual source identified.
- Variance exists in the extent of information contained in GILS records and their display (see Appendix H for two example GILS records that represent this variance).
- The service seemed qualitatively and quantitatively unpredictable and/or uneven.

While a majority of the users reported that they would *use GILS to locate government information in the future*, there were enough concerns and criticisms from the users to indicate that they consider GILS an unlikely source to help them identify and locate government information.

If users know of GILS, they make little use of it. When they do use GILS, they find it hard to use at best and inexplicable and frustrating at worst. Even agency staff involved in GILS implementations acknowledge that GILS is "user-unfriendly." Agency staff linked

the poor user reception of GILS to difficulties inherent in the search and retrieval system, the lack of full-text information, the limited direct links to the resource when discovered through a GILS record, and deficiencies in marketing GILS.

Users interact with specific implementations of GILS. While they may not recognize the elegance of the decentralized, distributed architecture, the construct of metadata for discovering resources, and the necessity of a robust information retrieval protocol, they do provide specific assessments of systems implementing the architecture and standards. Their assessments provide GILS designers and implementors with actual user requirements for what *users* want in a locator service.

GILS is in competition with agency Web servers. A participant in the online user assessment of GILS volunteered during the debriefing, in a positive, enthusiastic voice: "I *always* start with the agency homepage, and I find what I need about 40% of the time." The data from users indicate that the Web has had a dramatic effect on user expectations when locating and accessing networked resources. Users in the study's scripted online assessment continued to expect access to full-text of documents or access to services described by GILS records *even after they had spent time searching and were exposed to the construct of GILS as a locator*. From a user perspective, what GILS records describe is unclear and confusing.

4.4.8. FINDING: GILS Record Display Varies Widely and Is Criticized by Users

Most agency staff and virtually all users commented on the need to improve the content and display of GILS records. There is still considerable discussion and debate about the need for and use of specific data elements and the degree to which those data elements should be presented to users. GILS records were described by one person as "user-ugly." Appendix H presents two actual GILS records that exemplify the variation users may encounter as a result of a GPO Access GILS search on <"social security" AND pensions>. These

records show variation in content, format, and display. See also Appendix E-2 for examples of 4 high-quality records from the sample use in the record content analysis.

In the scripted online user assessment (see Appendix E-3), users commented on a number of presentation problems with GILS. First, since developers bill GILS as a "government information locator service," the majority of users suggested that *all GILS records should look alike*. There was also agreement with the statement: *The quality of records I examined varied widely*. Users recognized, and were disconcerted by, formatting errors (e.g., a record that did not have line wrap). Finally, there was frustration with not knowing "what they were looking at" on the screen or "what to do with the record." Investigators interpreted these comments to mean that users were not achieving an intellectual comparison between GILS and, for example, a record in a traditional or online library card catalog or a results list from a Web search engine.

The specifications for the GILS Core elements do *not* limit agencies in making improvements in the presentation of GILS records. Some study participants thought a "GILS-Lite" for presentation purposes is appropriate. A GILS-Lite record would offer the user a scaled down or reduced content record in an easier to read and use format. Additional research could determine the best or most useful collections of GILS record data elements to present to different users. Most of the agency GILS implementors, however, were unaware of how Z39.50 (the information retrieval protocol required by the *FIPS Pub. 192*) can provide different views of the record. The GILS Profile specifies several groupings of data elements to form "views" of the GILS record. But most implementations currently present the user with the entire GILS record.

4.4.9. FINDING: User Orientation and Instruction Is Inadequate

During the course of the study, the investigators found some agency online guides that provided basic introductory information to *their* GILS, but not a guide or manual that describes the GILS as a government-wide service and how best to use it, how best to conduct searches, and what kind of information and

output can be expected. Generally, training manuals and guides to assist users in their use of GILS are inadequate or non-existent.

GPO's manual *Helpful Hints for Searching Federal Databases Online via GPO Access* (March, 1996) is an example of the kind of guide that would be extremely useful for users to better exploit the GILS databases. The lack of training manuals, guides, or other such educational matter is part of the GILS marketing and visibility problem. The study finds that the lack of adequate user guides and related training material probably contributes to low use of GILS as well as frustration by those who do use GILS.

Agency officials, librarians, student users in the online assessment, and others contacted during the study gave low marks for the overall usefulness of GILS as a tool for identifying and accessing government information they needed. This is, in part, because they do not understand that GILS records were intended to describe metadata, *not* individual source documents. Confounding this is the occurrence of GILS records describing individual publications. Also of interest is the number of GILS implementors, GILS policymakers, and others who are involved in the actual development of GILS who are unfamiliar with its operation and use it infrequently, if at all. Thus, there is likely to be contradictory, confusing, or erroneous information disseminated about GILS.

4.4.10. RECOMMENDATION: Continuously Evaluate GILS Policies and Standards Against Emerging Technologies, Especially the Web

The emergence of the Web and its embrace by many Federal agencies for presenting information to the public have generated questions as to the role of GILS now that "we have the Web." At the time of GILS development, the Web was only minimally implemented. Given the near ubiquity of Web implementations by Federal agencies, a refocused GILS effort must determine how it can be integrated and evolve with the Web, as well as other emerging technologies (e.g. "push" technologies and natural language retrieval

systems). Refocusing GILS to support networked information discovery and retrieval may assist in that goal.

Many study participants acknowledged that the structured metadata record developed for GILS may be its lasting contribution. GILS should build on this success. Metadata can assist in the discovery and access of information in the networked environment. Standardized metadata is also independent of platforms and applications. Thus, the investment in GILS metadata should not be lost as the GILS evolves. GILS implementors will need to monitor ongoing metadata developments such as work on the Dublin Core and others.

New mechanisms for automatic indexing of networked information resources as envisioned by the Advanced Search Facility (ASF) deserve close attention. The ASF will provide an efficient means of gathering and indexing Federal information that goes beyond what current Web search engines offer. The complexity and difficulty of distributed search and retrieval of digital information cannot be underestimated. Networked information discovery and retrieval is still in its infancy and many issues and challenges remain (Lynch, 1995; Lynch, et al., 1995). GILS policymakers and implementors must have one eye focused on the future and the emerging technologies, and they must have the other eye focused on current citizens' needs for discovering and accessing information. The investigators think that effort expended in creating metadata records that support discovery and access will show a return on investment—library cataloging is a case in point. Technological solutions may assist in connecting users with government information, but the solutions must be workable and implementable.

The Web is a powerful existing technology for publishing and providing access to digital information. Its principle appeal is the hypertext linking within and between networked information resources to assist users in browsing and navigating full-text documents and how it enables user interaction with databases and online service. The Web's ability, however, to support networked information discovery and retrieval is limited. Existing Web search engines, while powerful, do not provide users with control and precision in searching

across Internet resources. Metadata, in the form of GILS records, can be used to enhance the discovery and retrieval of networked objects. Databases of GILS records can be a source for users to discover the existence of government information. Moreover, information in GILS records provide information not necessarily available to Web search engines for categories of information resources such as databases. Given the recommendation to limit a refocused GILS to online, network accessible agency resources, users can perform searches against GILS records using Z39.50 and then be linked to actual resources (e.g., full-text documents, other electronic resources and services). The investigators recommend that the next phase of GILS effort should strengthen the metadata functions and Z39.50 search and retrieval functions while continuing to explore and research integration with the Web and other emerging technologies.

4.4.11. RECOMMENDATION: Specify Resource Types and Aggregation Levels

To optimize the utility of an information system, a user needs knowledge of what information can be expected to be found in that system. In a library catalog, users can expect to find entries that describe items in a particular library's collection. An understanding of the unit of analysis (i.e., the granularity) of the items described in the catalog assists in its use. Catalogs usually represent a discrete item (e.g., one book) as the unit of analysis. Users have become accustomed to catalog entries representing books, as well as the scope and functions of the catalog. If users require representations of other units of analysis, they will often use other finding aids (e.g., indexes for journals to identify specific articles within a journal). For GILS to be a reliable and understandable aid in discovery, identification, and access to government information, users need to have a clear understanding of what information resources it includes and the unit of analysis for describing the resource.

The range of resource types and their granularity described in GILS is problematic. The original

vision of GILS intended GILS records to represent information resources such as existing locators, which might exist as a single publication or system, as well as aggregating resources not previously gathered or described as a collection (e.g., a set of press releases). The issues surrounding the granularity and aggregation of records and resources are complex, possibly more so because of the electronic nature of some of the resources.

From a user perspective, the issue of granularity and aggregation has several aspects. First, what can the user expect to be described by a GILS record? OMB 95-01 policy prescribes the description of three classes of information resources: automated information systems, locators, and Privacy Act system of records. These, however, are not necessarily mutually exclusive classes since a locator might be cast in the form of an automated information system. The actual practice of the agencies that are creating GILS records reflects the description of classes of resources beyond the three prescribed by policy; this was clearly evident from the record content analysis. If there are too many units of analysis being used, it is difficult for the user to know whether GILS will be useful for specific information needs. For example, can a user expect a GILS record to describe an individual document? Will it be a document that is in fact an index or locator, which the user examines to locate an individual document of interest? An understanding of the nature and scope of the refocused GILS equips users in information discovery. Users will need some understanding of the types of resources that might be discoverable through GILS. Further, the GILS records themselves should clearly identify the type of resource described in terms users can understand.

Throughout the study, some individuals stated that GILS needs to get users to the "real" or "actual" information. The implication of "real" or "actual" is that simply having a GILS record that describes a resource is not enough (although it can be easily claimed that just as a library catalog entry contains "real" or "actual" information, GILS records themselves are informative). Leaving a user with only a pointer is not sufficient, these people argue, especially if the resource itself is in electronic form.

A user perspective could argue in terms of the "distance" the user is from a resource that addresses or answers his/her information need. For most users, a GILS record is more useful if one can electronically link directly to the information object. For example OMB's *The Budget of the United States* GILS record describes a specific document, and, with the link provided, the user can retrieve and (via GPO Access) even search the digital version of that document.

An information object described by a GILS record, however, may be an online "locator" that the user would, in turn, search for desired information resource. An example of this would be GPO's online *Monthly Catalog* GILS record. GPO has a GILS record for the *Monthly Catalog*. In response to a user's search, the user may be presented with a GILS record for the *Monthly Catalog*. To continue the search for information pertinent to the information need, the user is required to do at least one more search—this time searching the *Monthly Catalog* to discover a citation for a specific resource. Although there may be no GILS record for the item described in the *Monthly Catalog*, the user is able to discover the item (and access the resource assuming it is in digital form and hotlinked from the *Monthly Catalog* citation).

These two cases of searching GILS illustrate how a user can move directly to a resource pertinent to an information need via GILS, or in the latter case, the user first uses GILS to identify another locator (e.g., the *Monthly Catalog*), and then conducts additional information retrieval transactions outside of GILS to find the desired information. One can discuss this in terms of "closeness" or "distance" from information objects, as well as traversing different "information spaces" to get to pertinent information resources.

The Web has been a conditioning force for Internet users. They have become accustomed to the experience of making several "clicks" and having at their disposal the "real" information (e.g., the full-text of a document, access to an online system). A refocused GILS with a more limited scope and coverage can support this type of information access, with the two examples give

above offering model approaches to providing this networked access.

To help users understand their "distance" from a resource described in GILS and the nature of the aggregation, an existing GILS data element, Resource Description, could contain a controlled value such as one from the list developed during the study's record content analysis:

- Subject matter database
- Publication
- Miscellaneous documents in ad hoc collection
- Agency homepage
- Organization
- Form
- Administrative catalog
- Bibliographic database
- Publications catalog
- System of systems
- Program
- Job line.

This list can be refined and developed so that a comprehensive list of GILS-described resources is available. As a part of the search results, where the user sees a brief form of the GILS record, the user could be presented with the resource type description along with a title and selected other GILS data elements. The brief form of the record should offer the user enough information to characterize the resource and enable the user to determine whether a particular resource described by a GILS record would be useful.

Further, a brief form of the GILS record should indicate whether the resource is network accessible, and by what means. While current GILS records occasionally include this information (i.e., by Available Linkage), users must read through many elements in the GILS record to discover it.

The investigators recommend that GILS policymakers and implementors should specify and define resource types to be described in the refocused GILS initiative. This determination should be informed by users' expectation to reach the full-text of a resource or link to another electronic resource.

Implementors should highlight the type of resource described by a GILS record and its network accessibility to assist users in making relevance judgments and accessing the needed information.

4.4.12. RECOMMENDATION: Enforce Consistent Use of Metadata That Are Empirically Demonstrated to Enhance Networked Information Discovery and Retrieval

The investigators encountered many comments related to the content requirements for GILS records and questions about the utility and benefit of the information included in GILS records. Based on these comments, and a refocused scope for Federal GILS implementations, there is a need to review the data elements as used in agency GILS implementations with the goal of optimizing them to support the discovery, identification, and access of government information.

Information organization begins with a selection and filtering process and a distillation of essential features from each information object (Hsieh-Yee, 1996). A point of contention becomes immediately obvious: what is valuable or essential? Those who seek to make "resource discovery and retrieval" possible in the networked environment must determine which information resources are worth describing, a significant initial step. But a second set of decisions may be even more difficult—those concerning the salient features of the information resources that need to be represented and described in a record.

The data elements for GILS records had their genesis in an interagency working group. Different stakeholders within that group identified data elements necessary to support specific functions. Record creators need to collect or capture the information to provide content for data elements, recognizing that each bit of information included in a GILS record has an associated cost. Which are the highest value pieces of information that could be included? How much information should be contained in a GILS record? These are not easy questions to answer, especially given the

diffuse goals, purposes, and expectations of GILS discussed earlier.

GILS is a pioneering effort in what has become a major research and development activity (i.e., determining appropriate metadata schemes for networked information discovery and retrieval). The community of interest that defined the initial set of metadata (i.e., government agency staff) had particular requirements for GILS, and these requirements were codified in the appendix of the GILS Profile that identifies and defines the GILS elements. A key question at this point is: what are the salient features of an information resource that need to be represented in a GILS record to support users discovering, identifying, and accessing U.S. Federal government information? An associated question is: do different classes of resource types need different groupings of metadata elements (e.g., if one is representing a document rather than a database rather than a Web site).

The work on the Dublin Metadata Element Set could inform a review of the data elements for a refocused GILS. The goal of the Dublin Metadata Element Set is to devise a simple and minimal metadata scheme to provide descriptions of one class of networked information resources (i.e., document-like objects) for their discovery and retrieval. The 15 elements of that metadata scheme may be sufficient for the revised purpose of a modified GILS—namely the discovery and access of government resources. This approach should be explored in the context of reviewing the existing 67 mandatory and optional GILS elements.

Including metadata elements in GILS records that support objectives other than the public's discovery, identification, and access of information (e.g., IRM and records management) confounded GILS implementation. In the next stage of GILS development, the overriding criteria for determining mandatory metadata should be driven by the newly articulated purpose and goals of a refocused GILS initiative and the uses to which the metadata records will be put. Policymakers may find a review of the development of Federal Geographic Data Committee (FGDC) metadata helpful; that community identified four criteria for inclusion of specific data elements (Mangan, 1995):

- **Availability:** information needed to determine what data exist for a given geographic area
- **Fitness-for-use:** information needed to determine if a dataset meets a specific need
- **Access:** information needed to acquire an identified dataset
- **Transfer:** information needed to process and use a dataset.

Regardless of the criteria for determining GILS data elements, a formal process is needed for discussing and identifying the U.S. Federal requirements for data elements to support users' needs to discover and access government information resources. The formal process requires identifying an agency or interagency body as the official forum for discussion of U.S. Federal GILS specifications as well as acting as a "steward" of the GILS data elements as used in U.S. Federal implementations (e.g., developing guidelines for record creation, providing assistance in using the data elements, etc.).

The current process for revising GILS data elements is under the jurisdiction of the GILS SIG, with discussion on the elements occurring in monthly meetings of the GILS SIG and through the GILS Forum, an online discussion group established in 1994. According to the GILS SIG statement of purpose, All recommendations developed at the periodic meetings will be distributed via the listserver [i.e., the GILS Forum] for comment and additional discussion prior to becoming final" (see Appendix A-6). The Forum is open to anyone with access to an Internet email account and is not limited to U.S. Federal government agency staff and associated stakeholder communities. The GILS Profile is a general purpose profile for describing and locating information, not exclusively government information. The U.S. Federal implementation of GILS has specific requirements, and it is appropriate that a formal body (agency or interagency) be authorized with the responsibility for stewardship of the data elements scheme in the next stage of GILS development. Such a forum, however, must coordinate efforts with other

government agencies that are promoting one or more metadata schemes (e.g., NARA's records management data elements, FGDC content standard, etc.).

The investigators recommend that metadata elements should be reviewed within the context of the revised and more focused purpose for GILS, namely discovery, identification, and access of government information. Data elements should be included/excluded in the metadata scheme based on the extent to which they demonstrate support of enhanced discovery and access of government resources. An analysis of the cost/benefit of the current data elements compared with their capability to support of the purpose of a refocused GILS should be done. GILS metadata development should also take into account activities of other major groups that are developing and evolving metadata schemes and the evolving technology that supports distributed search and retrieval.

Either a single agency or an interagency group (e.g., a GILS Committee of the CIO Council) should be charged specifically with the review, development, maintenance, and revision of GILS data elements as used in U.S. Federal GILS implementation. GILS policy should identify the body responsible and direct it to prepare specific written and easily available procedures and criteria for maintaining and revising the GILS metadata elements. The resulting process will provide agency implementors to determine new elements or modifications to existing elements based on the requirements of a refocused GILS. After Federal implementors identify their requirements and proposals, these can be forwarded to the GILS SIG, which has authority for maintaining the GILS Profile. This process recognizes that U.S. Federal implementors may have requirements different from other communities that use the GILS Profile.

4.4.13. RECOMMENDATION: Improve Presentation of Metadata

Users, whether agency staff, librarians, public users, or others, noted problems with the presentation of GILS records. They remarked about records containing too much information, or not the right information, difficulty in understanding what the

GILS records described, and the unpredictability of element inclusion (e.g., use of nonmandatory elements or locally-defined elements).

The development of the GILS Profile acknowledged that different user groups might need different views on the GILS data elements. Although one might question whether the Profile defined appropriate and adequate data elements in the first place, the issue of presenting GILS data elements in the records is quite separate.

Policymakers chose Z39.50 as the information retrieval protocol to support GILS because of its functionality in providing a uniform interface to different information servers and their associated databases. It also allows Z39.50 clients to request different views of the database record (e.g., a GILS record). Thus, it separates searching records (enhanced by the number of structured access points available) from presenting the records (which can be customized by implementors). For example, the GILS Profile identified several views of the record, where each view presented different amounts of data to the user. The key question remains: what is the appropriate information to present to users, and at what stage in the search/retrieval process?

Current GILS implementations using Web-based interfaces usually present, in response to a search, a result set of "hits" (i.e., pointers to GILS records that meet the criteria of the search). Users of GILS are first presented with the list of "hits" in the result set, and those hits are generally represented by only the title and a relevance score. When users select a GILS record from the result set, most agency GILS implementations display a view of the complete GILS record. The question is: is the complete GILS record the appropriate or only view of the record to present to the user? As discussed earlier, study participants proposed a "GILS-Lite" record that would present a briefer view of the entire record. Such views can be accomplished using Z39.50.

Experience from the Cyberstacks project at Iowa State University (McKiernan, 1996a) suggests that record creators need not "delineate all relevant elements in describing a resource," but rather

should "characterize the resource sufficiently so that users can judge its potential usefulness" (McKiernan, 1996b). McKiernan recommends that users need only an appropriate characterization to determine whether an item is potentially relevant and deserves a closer look.

Relevance and selection judgments by users comprise a complex process (Barry, 1994). Agencies need to experiment with providing different views of GILS records to their users to determine which views are appropriate at different stages of the information retrieval process. The investigators recommend that agencies should remember the purpose of a refocused GILS and experiment with presenting users with different groupings of data elements. Such experiments should be evaluated closely, and the experiments themselves should be informed by recent and ongoing research in user relevance judgments, as well as human computer interface design (Schneiderman & Croft, 1997).

An interagency effort should be mounted to address issues of presentation and use of metadata records (e.g., when they should be presented to the user, when should the use of GILS be transparent to the users, which data elements to present, etc.). Speed and ease of finding the information (e.g., identify a maximum number of "clicks" to get the user to the GILS record and the described information resource) should combine with readability, consistency, layout, and other presentation features of the record to optimize information discovery and retrieval. The GILS pilot program offers a venue for the development and testing of two or more Z39.50 clients that support the function of element selection and processing for customized display to users.

4.4.14. RECOMMENDATION: Develop Policy and Procedures for Record Maintenance

Although many agencies have created GILS records, the maintenance of those records appears to be less well-supported. As noted earlier, agencies that see no benefits from GILS have little or no incentive for continuing to create more records or to maintain the records they have created. Keeping metadata records current and accurate should become part of the day-

to-day fabric of agency information resources management activities.

GILS will not endure unless agency staff consistently maintain GILS records. This is especially important because of time-sensitive data included in the records. For example, a set of elements in current GILS records hold information about the point of contact, including contact names, telephone numbers, and email addresses. Further, where a GILS record contains a pointer or link in the form of a Uniform Resource Locator (URL) from the record to the described resource, staff must ensure that the link is still operable.

In the decentralized environment for agency GILS record creation, the investigators recommend intra- and inter-agency efforts at devising written policy and procedures for record maintenance. Such policy and procedures should address the varying levels of networked infrastructure in agencies as well as other factors such as intra-agency cooperation from offices of primary interest in record maintenance and updating. There will be the need for mechanisms to automatically remind the record creators that their records need review and/or updating. Different types of agency resources may be more subject to change than others, and thus need more frequent maintenance. Software that tracks the date of last modification of a GILS record could trigger an alert (e.g., in the form of an email message or utilizing "push technology") record creators to review their records and update them if necessary.

The issue of record maintenance must be addressed since the degradation of the currency of GILS records will hinder access to government information. A GILS pilot program offers an opportunity for fine-tuning the policy, procedures, and software for maintaining GILS records.

4.4.15. RECOMMENDATION: Promote Interagency Cooperation and Use of GILS for One-Stop Shopping Functionality

Several agency and interagency initiatives use the Web to provide one-stop shopping to collections

of government information resources and services. These include:

- Business Advisor, the one-stop electronic link to government for business
<http://www.business.gov/>
- Federal Statistics Initiative
<http://www.fedstats.gov>
- WINGS, Web Interactive Network of Government Services
[http://www.wings.usps.gov/](http://www.wings.usps.gov)
- Commonly Requested Services
<http://www.whitehouse.gov/WH/Services/>
- National Environmental Data Index (NEDI)
<http://www.nedi.gov>.

One can think of users needing government information about particular topics. Often these information needs are not formulated, nor need they be, in terms of "what agency should I contact to get the information I need?" Rather, users may think in terms of "where can I find government statistics on unemployment rates and their impact on welfare requirements?" In the latter case, the collection of resources from various agencies (e.g., Department of Labor or Department of Health and Human Services) in a one-stop shopping scenario is more effective than presenting information according to the missions of government departments, agencies, and bureaus.

Implementors structured the Federal GILS initiative along agency lines, but this basis of agency locators does not preclude interagency one-stop shopping scenarios. GILS provides a mechanism for agency resources to be identified and described. For any particular topic area (e.g., environment, energy, etc.), the relevant GILS records could be gathered and placed in a database for user searching. This, when combined with the Web-based initiatives listed above, provides users with several means of access (e.g., browsing and free-text searching of a Web site, and searchable GILS records for identifying specific resources of interest).

OMB Bulletin 95-01 language reflects this approach:

Interagency committees which promote access to and use of Federal information are encouraged to

coordinate the efforts of their participating agencies in developing their respective GILS inventories and interagency topical locators when appropriate to their respective missions.

During the evaluation study, the investigators saw some evidence of such cooperation vis-a-vis GILS. For example, EPA participates in NEDI, which "contains" EPA's GILS records. Given the example of NEDI, it is likely that such coordination may occur only if agencies see they are addressing real information needs and are doing so to accomplish their missions. Identifying exemplary interagency cooperation that builds upon individual agency GILS efforts to serve as models and offering incentives could assist in the development of one-stop shopping.

In addition to cooperative efforts that provide a single point of access to collections of resources thematically or topically related, users also identified a desire for one-stop shopping for searching for government information. This part of the vision of GILS has yet to be realized, in part because of the very real difficulty and complexity of conducting distributed search and retrieval. GPO is experimenting with cross-agency GILS searching, and this effort should be applauded as well as evaluated. Such efforts need to assess if searching and retrieval performed under this configuration increases user satisfaction with results. The Advanced Search Facility (ASF) also may offer a technology solution to cross-agency searching. Based on the information gathered during the study about ASF, GILS implementors should follow its development closely.

The investigators recommend that GILS policy promote interagency cooperation and provide incentives to realize one-stop shopping for government information. A refocused GILS should have as its goal the support of government-wide searching for information. GILS should provide a means for users to discover and access information on a *government-wide* basis. Distributed searching across all agencies' information resources (e.g., by searching across agencies' GILS databases) provides one approach to one-stop shopping. Interagency cooperative

efforts should be encouraged to develop collection of government resources, and the refocused GILS initiative can explore how GILS can support such efforts.

4.5. OPPORTUNITY: RESOLVE GILS RELATIONSHIPS WITH OTHER INFORMATION HANDLING FUNCTIONS

Agencies' management of their information resources involves many different information handling functions (e.g., providing public access, inventorying, records management, etc.). The establishment of GILS added yet another function. This opportunity addresses findings and recommendations related to the role of GILS vis-a-vis other agency information handling processes.

One of the challenges in the next phase of GILS development will be to resolve how GILS fits with and can be integrated into these processes. Table 4-8 summarizes the findings and recommendations for this opportunity.

4.5.1. FINDING: GILS Does Not Support Records Management Activities

OMB Bulletin 95-01 identified a records management component for the Federal GILS initiative. GILS designers and researchers, however, did not consider GILS as a tool to support records management (Moen & McClure, 1994a; Information Infrastructure Task Force, 1994). While GILS policy considered public access and records management mutually supportive, the study identified significant problems with using GILS as a records management tool as outlined in OMB Bulletin 95-01. Identifying a records management component for the Federal GILS initiative led many agencies to delegate GILS implementation to records managers. While records managers have responsibilities related to identifying information resources for scheduling and archival purposes, it is not clear that records managers were in an appropriate position to recognize the broader possibilities and benefits for public access and IRM that the GILS initiative could support.

Table 4-8
Resolve GILS Relationships with Other
Information Handling Functions

OPPORTUNITY: RESOLVE GILS RELATIONSHIPS WITH OTHER INFORMATION HANDLING FUNCTIONS	
Findings	Sources of Evidence*
4.5.1. GILS Does Not Support Records Management Activities	FG, KP, SV
4.5.2. GILS Relationship with Agencies' Inventories of Information Resources Is Not Clear	CA, FG, SV
4.5.3. GILS Relationship with FOIA and EFOLA Is Unclear	FG, SV
Recommendations	
4.5.4. Uncouple the Refocused GILS—as an Information Discovery and Access Service—from Records Management	
4.5.5. Derive GILS Metadata from Other Information Handling Processes	

* CA=content analysis of GILS records; FG=focus group sessions; KP=interviews with key participants; LA=log analyses of Web servers; SU=survey conducted at the 1996 GILS Conference; SV=site visits to selected agencies; US=scripted online user assessments of GILS

From a records management perspective, specifically in terms of records scheduling and the information needed for scheduling records, much of the information GILS records describe is not organized in such a way as to be useful. Records managers schedule records in series and do not manage individual publications or documents. They schedule publications (e.g., information dissemination products) as part of a series, often a series that describes the agency's information dissemination products as a whole. For automated information systems, records managers schedule not only the system itself but its inputs and outputs. As noted in the NARA booklet on managing electronic records, "It is also essential to emphasize that all components of electronic information systems are records: inputs, outputs, digital data stored in a variety of ways, and the related documentation" and each of these different record components of an information system may be on different retention schedules, etc. (National Archives and Records Administration, 1990, p. 5). these are not currently described by GILS records.

Some agencies create GILS records for individual publications such as the Internal Revenue Service's *Catalog of Federal Tax Forms, Form Letters, Computer Generated Letters and Notices* or the

Department of State's pamphlet *Americans Abroad: What You Should Know Before You Go*. Other agencies have one GILS record for all publications such as the Federal Emergency Management Agency (FEMA) GILS record for its *FEMA Publications Catalog*. The variety of aggregates and information types that GILS records describe makes these records ineffective for records management purposes.

The GILS data elements do not support records management since they do not account for important information such as record retention and disposition in ways that serve records managers. OMB Bulletin 95-01 directed NARA to:

Cooperate with agencies to reduce reporting burdens and facilitate scheduling of records by accepting GILS data entries when they provide the information required on Standard Form 115, Request for Records Disposition Authority.

Yet the records managers interviewed during the study stated flatly that GILS records were not adequate for records management purposes. More importantly, records managers expressed little enthusiasm to "enhance" GILS data elements to carry such information, in part because of the mismatch of records management practices (e.g., scheduling records in

series) and the types of resources described in GILS and related granularity/aggregation issues.

OMB Bulletin 95-01 also included guidelines to agencies in terms of their responsibilities for using GILS in records management functions:

By December 31, 1996 [all Federal agencies will] submit to the Archivist a request for disposition authority proposing schedules for unscheduled records in the information resources described in the GILS Core locator records. The agency should also advise the Archivist if it believes any information resource described in the GILS Core locator records has sufficient historical or other value to warrant continued preservation after the information is no longer needed in the agency. [Section 4 (4)].

Policymakers envisioned GILS as a mechanism to discover and identify agency records in need of scheduling. Several study participants mentioned that in practice this did happen occasionally.

But in considering GILS as a tool for NARA to use in monitoring agency resources that had not yet been scheduled, NARA representatives said that GILS is not comprehensive and would not be reliable as the only tool for them to use. GILS *may be* useful as a finding tool to uncover material that should be scheduled and to enhance the thoroughness of agency records management. But in fact, the investigators identified only a very limited use of GILS for these purposes.

NARA representatives also told the investigators that NARA received few SF 115s from only limited number of agencies because of the OMB Bulletin 95-01 requirement. NARA detected no significant increase in the number of scheduling requests attributable to GILS. They concluded that GILS was not having any major impact on scheduling agency records. While the OMB Bulletin viewed GILS records as carrying data that would make submitting the SF 115 redundant, the SF 115 is the *legal* instrument used in the scheduling process. Further, the SF 115 is only one part of the scheduling process, a process that includes

authorizing signatures, etc. GILS policy on records management seems not to have recognized this fact.

Findings from the study suggest that GILS is not equipped to improve government-wide records management activities and responsibilities. A government-wide system for records management is needed. The current GILS, however, is not the system to accomplish it. Curiously, agencies perceived or suspected a connection forged between GILS and records management as something devised by GILS creators in collaboration with NARA to provide political support for the records management function in Federal IRM and not something that arises out of a natural affinity between GILS and records management.

4.5.2. FINDING: GILS Relationship with Agencies' Inventories of Information Resources Is Not Clear

Agencies are required by OMB Circular A-130 to develop and maintain inventories of their information resources. A previous study by the investigators (McClure, Ryan & Moen, 1992) identified agency locators, but concluded that agencies did not have in place comprehensive locators to their information resources. One of the assumptions of OMB Bulletin 95-01 was that such agency locators did exist, and that creating GILS records describing these locators would not be a major burden on the agencies. As noted in Section 4.3.9., even though the policy required agencies to inventory their resources, no clear guidelines and prescriptions emerged to guide agencies in how those inventories could become useful network-accessible locators, which in turn could be described by GILS records.

In discussion with study participants, representatives from the small agencies noted that oftentimes such locators did not exist, and they expressed strong feelings of anger and frustration against the assumption that inventories were in place. This false assumption allowed OMB to assume that GILS record creation would be relatively effortless. Given current realities, the smaller agencies are unlikely to participate in future GILS activities without significant changes in the GILS initiative.

Other study participants acknowledged that implementing GILS forced some of them to accomplish some inventorying of their resources. Both agency staff and other GILS stakeholders said that this was a positive byproduct of GILS.

It is unclear whether GILS should be seen as the tool to gain agency compliance with developing information inventories. In the discussion of GILS and records management, study participants noted that because GILS does not provide a comprehensive list of agencies' resources, its utility for that aspect of records management (i.e., discovering what resources might exist that are in need of scheduling) is limited. One can conclude that GILS is not moving agencies, especially the smaller ones, to a comprehensive coverage of their resources and has not become an inventory of agency resources.

4.5.3. FINDING: GILS Relationship with FOIA and EFOIA Is Unclear

The passage of the *Electronic Freedom of Information Act Amendments* (EFOIA) (P.L. 104-231) in Fall 1996 immediately preceded the 1996 GILS Conference. At that conference, Sally Katzen (1996) of OMB stated:

Second, GILS could become the "killer application" for agencies to use in implementing the provisions of the new *Electronic Freedom of Information Act Amendments of 1996*, which President Clinton signed into law just last month, and which contemplates a more proactive approach to agency identification and access to important records.

Her statement had considerable impact on the audience and was mentioned to the investigators repeatedly during later data collection activities.

One set of viewpoints identified by the study regarding GILS and EFOIA is, indeed, *what if* GILS could become the "killer application" and provide the means by which agencies could implement various record keeping provisions of EFOIA? But upon further discussion, specific

strategies for accomplishing this objective, how the GILS records data elements would need to be changed, and the level of effort to "shoe-horn" EFOIA provisions into the GILS concept were unclear at best.

Reactions to Katzen's statement exemplified the multiplicity of understandings (and misunderstandings) of GILS intent and potential. As one person commented to the investigators, "it's just another unfunded mandate by OMB that hasn't a clue as to the level of effort and resources needed to make it happen." At one focus group session of agency records managers, participants laughed at the idea that GILS, as presently constituted, could begin to address the EFOIA functionality that Katzen mentioned.

The timing of Katzen's statement provided a catalyst for this discussion to occur during data collection activities. To some extent the debate about the role of GILS in EFOIA is a microcosm of the larger GILS assessment: What is GILS' purpose versus potential purposes? How will GILS initiatives be funded and implemented at the agency level? What changes in GILS record content will be needed? Who will provide the leadership to develop this "killer application?" Overall, study participants found this "opportunity," as suggested by Katzen, to be but another task for which they had no time, staff, or other resources to address.

The policy review in Chapter 2 discussed EFOIA and a recent memorandum from OMB that links GILS and EFOIA (Office of Management and Budget, 1997a). Yet the guidance in the memorandum (i.e., agencies should establish a GILS "presence" to address requirements of EFOIA) lacks precision and begs the question as to how—specifically—GILS can assist in handling EFOIA requests.

Resolving the issues of integrating GILS and EFOIA is beyond the scope of this study and requires additional study before a recommendation could be made. A research effort could examine a range of FOIA requests to determine what information a GILS record would need to contain to assist the user in identifying the object of the sampled FOIA requests. Proposal for using GILS to support EFOIA will require careful assessment and study to determine what, if any, real linkage can be made between GILS and EFOIA. The

GILS pilot program would provide one opportunity for such study.

4.5.4. RECOMMENDATION: Uncouple the Refocused GILS—as an Information Discovery and Access Service—from Records Management

The findings in the study are unequivocal about the lack of utility for records management provided by GILS in its current implementation. While there may be some logical connection between a locator service and the records management responsibilities of agencies, the U.S. Federal implementation of GILS does not justify GILS as a records management tool.

The discussion of findings above offered reasons why GILS does not support and is not a suitable mechanism for records management (e.g., granularity of records, availability of data elements to carry records management information, etc.). Although there was some evidence that GILS records could be used for identifying resources that need to be scheduled, GILS is limited in utility in this records management function as well since currently GILS cannot be relied on to represent comprehensively the resources of an agency. One potential use of GILS that intersects with records management would be to require agencies to create GILS records that describe and point to agencies' records schedules. Ideally, the schedules themselves should be network-accessible, and users could discover and locate the schedules, and then use the schedules to identify agencies' information resources.

Three options are possible relative to GILS and records management:

- Make no changes to GILS related to records management and assume that agencies will try to use GILS to some extent for their records management activities.
- Enhance GILS by adding additional data elements and other specifications to help

creators of GILS records provide the information and describe at the appropriate level of granularity to serve records management goals.

- Sever GILS from records management activities.

While all three of these represent possible directions, the investigators recommend the third option. The evidence from the study was substantial—from the perspective of records managers and NARA—that GILS is not suitable for records management, and in particular for supporting records management processes such as scheduling and communicating scheduling information in lieu of the SF-115. Further, there was no agreement on how GILS could be enhanced or changed to make it a usable tool for records management, nor that the effort in doing so was warranted.

Uncoupling the Federal GILS initiative and records management will bring several issues to the fore. If policymakers designed GILS as a means by which the "electronic records management" problem could be solved, the study concludes that GILS is not the solution. In fact, the term "electronic records management" can refer to the management of electronic records (simply applying records management activities to resources that are in electronic or digital form) or to using information technology to support processes involved in records management such as electronic submission of SF-115, digital signatures, etc. GILS policy appears to have emphasized the latter aspect, and GILS does not support that aspect of electronic records management.

Senior staff at NARA are aware that GILS is not serving records management purposes, and also realize that government-wide electronic records management needs a solution. Although GILS policy should not address records management issues, policymakers should expect NARA, in collaboration with the agencies, to develop a workable solution to government-wide electronic records management in a realistic timeframe (e.g., 16–24 months).

NARA should develop a formal program to implement records management processes and procedures that will allow agencies to submit

electronically requests for records scheduling and disposition authority (i.e., an electronic version of SF 115). NARA's program should be developed with input and advice from policymakers and agency officials, and the program should include specific objectives and time frames for monitoring its progress.

At the agency level, an impact of the uncoupling of GILS and records management raises the question of who in the agency will be responsible for GILS? Many agencies, especially the smaller agencies, delegated GILS responsibility to their records managers. This was likely due to the records management language in OMB Bulletin 95-01. If GILS does not play a role in records management, it is likely that agencies should and will identify non-records managers with responsibilities for agency GILS efforts. This raises an important question about who, in the agencies, are best positioned to assume the responsibilities for implementing a refocused GILS?

NARA's responsibilities per OMB Bulletin 95-01 for developing guidelines and providing training for GILS record creation will need review in the next phase of GILS. Such guidelines and training will be needed in the refocused GILS, and NARA brings appropriate expertise related to content standards and descriptive records. GILS planners must identify training and documentation as key areas for attention, and NARA (or possibly the cataloging expertise at the Library of Congress) could be a resource in the development and provision of training.

The investigators recommend that revised GILS policy should uncouple the discovery, identification, and access function of a refocused GILS from agencies' records management responsibilities. GILS will not and should not be used as a mechanism for solving a range of electronic records management problems. There is no apparent natural affinity between public-access NIDR and electronic records management, and both programs must be sufficiently mature before viable connections between them will be made.

4.5.5. RECOMMENDATION: Derive GILS Metadata from Other Information Handling Processes

An important aspect of a refocused GILS effort will be to identify how GILS can be and should be integrated into agency information handling processes. In particular, the refocused GILS effort should identify ways to prevent agency GILS activities from *dis-integration* with other information handling and dissemination processes. For many agencies, GILS implementation has been a standalone add-on, which weakens its benefits and buy-in. For the new effort to be successful, it must be integrated into other information handling processes.

One of the primary benefits of GILS to date is the creation of standardized, structured records for describing agency information resources. These metadata records are essential for the discovery and retrieval of information in the networked environment. One important area for development is to determine how GILS metadata can be captured automatically for each new information resource created by an agency.

The refocused GILS initiative must address how GILS metadata can be captured in the most effective and efficient way. Discussions with agency staff responsible for creating GILS records noted the significant level of effort to collect content for the records. The actual inputting of the GILS record is relatively trivial in terms of labor, but the collection of adequate, accurate, and appropriate information is extremely time consuming. Too often the people in charge of creating the records did not have the information available to them, and too often the record creators did not gain cooperation from agency staff who were primarily responsible for the information resources that needed to be described in GILS.

The *retrospective* character of GILS record creation is a problem. Agency resources exist; data must be collected retrospectively about those resources prior to the creation of the GILS record. While it is possible to do such *retrospective cataloging* of agency resources, the cost of creating GILS records for all agency information resources may far outweigh the benefit. Clear guidance is needed on

what existing resources agencies should describe in the refocused GILS.

Existing agency resources are just part of the government information universe GILS addresses. Each day, new agency resources are created and developed and added to the information universe. Since most current agency information resources begin as an electronic file in the information life cycle, electronic document management systems (EDMS) may contribute a solution to GILS record creation for at least some categories of new information resources. Metadata for document-like electronic resources can and should be captured at the point of creation, and EDMS provides mechanisms to do this. Agencies could use the captured metadata for creating item-level GILS records or could compile item-level metadata into agency locators, which in turn can be described by a GILS record.

The refocused GILS effort should examine the most efficient means for capturing basic metadata whenever a new information resource is created or initiated. For this to succeed, a refocused GILS cannot be a standalone system but rather GILS records need to be derived from metadata captured in the process of creating and managing an information resource through its life cycle. EDMS should be exploited by agencies to manage their electronic information resources (e.g., document version control, reduced duplication of effort, inventory reporting, etc.). The system can incorporate a module whereby metadata about the resource being created can be derived.

The metadata to be captured, and when, should be informed by the purposes the metadata serve. A refocused GILS should determine the appropriate metadata to support information discovery, identification, and access. Appropriate and accurate metadata can be more easily determined and assembled during the process of creating the information resource than after. As part of managing an information resource through its life cycle, agencies may need to capture metadata that serve purposes in addition to information discovery and access. The focus for the next phase of GILS, however, is to identify the metadata needed to serve purposes of a refocused

GILS and identify effective ways of capturing the metadata.

While discussions of electronic document management systems are outside the scope of this report, policymakers and agency information managers need to make the systematic management of electronic documents a priority policy area. To manage agency electronic resources systematically requires an understanding of an agency's information processes and flows (i.e., an architecture) and a focus on information life cycle management (Ambur, 1996). One component of the architecture will be the capture of appropriate data about electronic information resources, and the capture of GILS metadata information can be accomplished within such a scenario.

The investigators recommend that policymakers and implementors explore and assess various practices to integrate GILS into existing or emerging information handling processes and systems. Without integration, GILS may be subject to lack of attention as a standalone activity. A critical aspect of its integration will be in determining the best practices for capturing GILS metadata at the time of creation of new information resources, and EDMS can serve as one model for automatic capture of metadata. Based on comments by study participants, little increase in the number of GILS records is likely unless the process of capturing metadata is less labor intensive and more cost-effective.

While there are many reasons for the uneven character of agency GILS implementation (e.g., lack of tangible benefits, cost of creating records, lack of management support, etc.), an approach that ties GILS into other information handling processes may assist agencies in reconsidering the utility of GILS. If GILS activities are not integrated, and if metadata capture cannot be made less burdensome, even a refocused GILS effort may be threatened by current resistance to GILS—it will remain "one more thing" agency staff have to do separately from other activities.

4.6. OPPORTUNITY: INCREASE GILS AWARENESS

Except for a relatively small number of study participants who have been intimately involved in GILS activities and implementations, the study found the majority of participants lacking in basic understanding of what GILS was intended to be and how it was to function. Outside the "beltway," the investigators found minimal awareness that GILS existed, even among important user communities such as government documents librarians.

For any product or service to succeed, a program of promotion and education is necessary. The following findings and recommendations address the need for a refocused GILS to increase awareness about the service, but to do that the purpose and goals of GILS must be clarified so a coherent message can be delivered about the service. Table 4-9 summarizes the findings and recommendations for this opportunity.

4.6.1. FINDING: No Program for GILS Promotion and Education Exists

The study explored the extent to which a coordinated promotional effort for GILS exists, and who, specifically, was charged with responsibility for that effort. By and large the answer is that there has been no *government-wide* campaign effort for GILS.

A number of champions and spokespersons have come forward to talk about and support the GILS efforts. The Special Interest Group on the Government Information Locator Service (GILS SIG) developed and disseminated a one-page brochure describing GILS, but, according to one of the participants in that effort, gaining consensus on what the brochure should contain and how to state the purpose of GILS was a challenge and involved several months' discussion.

In addition, different spokespersons "marketed" GILS differently, leaving contradictory messages of why GILS was important and what GILS was intended to accomplish. One result of the lack of

"marketing the product" and "keeping on message" about GILS could be the varying expectations of GILS encountered in the course of the study. The absence of a central and coherent message allowed GILS to become "different things to different people."

NARA provided training sessions for GILS implementors, specifically for those involved in record creation and the use of the NARA *Guidelines*. Such training sessions, however, answered only one aspect of the education needed by agency staff to understand why GILS is important and how it can be used to improve public access and agency management of information resources. Efforts for systematic training for GILS users were minimal. The notable exception has been GPO's training of documents librarians on GPO Access, which now includes a hands-on session for GILS.

Many agency staff that participated in the study criticized the lack of government-wide or other systematic promotion of the Federal GILS initiative. Study participants remarked that no single and unified voice came forward in the past two years to market GILS. The lack of a program promoting GILS resulted in a low level of awareness and limited acceptance of and support for GILS.

Advocacy of GILS was difficult because of the confusion over its purposes—what it was supposed to offer, how it worked, what people could expect to find in it. This finding reveals the need first to define GILS and then develop a strategy for promoting it. Refocusing the GILS effort by identifying an understandable scope and function of GILS will be an important first step.

One site visit participant suggested that there was need for a public marketing campaign for educating people about GILS—something they saw had not been done to date. Another person suggested that "GILS should be promoted as 'this is how you find information about the government' and make it a central and first point of contact for finding government information or more general information about the government." Given the various problems with GILS, the absence of a promotional campaign is probably not significant. It could become a deciding factor, however, in the success of a refocused GILS.

Positive promotion of a refocused GILS can reap benefits within an agency. An agency site visit participant suggested that "many agencies do not grasp the significance or potential for GILS and an all-out marketing effort by OMB and NARA needs to address this. Such a dual marketing strategy would assist in getting more top management support that would, hopefully, filter down to the bureaus and department levels of agencies." Senior agency management needs to make the refocused GILS a priority if it is to be successful, and a promotional campaign directed at agency managers could be effective in garnering additional resources and commitment. Moreover, a GILS pilot program could demonstrate how GILS works and the benefits from using GILS. Thus, the program of promotion will serve to educate agency managers and staff as well as non-government users about GILS.

At an agency level, study participants identified a number of potential benefits from systematic promotion of Federal initiatives:

- Greater senior management buy-in
- Active demonstrations of the utility
- Extension of participation (e.g., creating and maintaining more GILS records)
- Enhancement of applications (e.g., data gathering and input for GILS records)
- Improved training to agency units
- Development of marketing tools such as brochures.

Systematic promotion is thus essential for GILS—from a policy perspective as well as management, implementation, and use perspectives. Such marketing has not happened to date in the GILS initiative, nor was any one agency charged specifically to develop and carry out such a marketing program. Lack of product marketing reflects an *if we build it, they will come* attitude.

4.6.2. FINDING: Potential User Communities Lack Familiarity with GILS

The study found very low visibility and limited knowledge about GILS outside a core group of champions, policymakers, and agency implementors. The survey distributed at the Fall 1996 GILS Conference asked respondents to rate their familiarity with GILS documents and policies. Of this group of people that could be considered knowledgeable or at least interested in GILS, less than 50% claimed familiarity with some of the basic GILS documents and specifications. (In contrast, a majority of respondents claimed familiarity with the World Wide Web.) Most problematic is the lack of familiarity by this selected group of people at the GILS Conference with three basic GILS documents: OMB Bulletin 95-01 (policy); NARA's record creation guidelines (for implementation activities); and *FIPS Pub. 192* (for technical specifications and guidance). Table 4-10 summarizes the responses (see also Tables E1-5 and E1-6 in Appendix E-1 that contain the complete survey results).

**Table 4-9
Increase GILS Awareness**

OPPORTUNITY: INCREASE GILS AWARENESS	
Findings	Sources of Evidence*
4.6.1. No Program for GILS Promotion and Education Exists	FG, SU, SV
4.6.2. Potential User Communities Lack Familiarity with GILS	FG, SU, SV, US
4.6.3. GILS Usage Is Limited	FG, LA, SU, SV, US
Recommendations	
4.6.4. Develop and Formalize GILS Promotion, Education, and Training Strategies	

* CA=content analysis of GILS records; FG=focus group sessions; KP=interviews with key participants; LA=log analyses of Web servers; SU=survey conducted at the 1996 GILS Conference; SV=site visits to selected agencies; US=scripted online user assessments of GILS

Table 4-10
Familiarity with GILS Documents/Policies

GILS Documents/Policies	Familiar		Neutral		Not Familiar		Blank		Total	
	N	%	N	%	N	%	N	%	N	%*
Federal GILS Policies	86	48	40	22	52	29	3	2	181	101
Agency's GILS Policies	81	45	24	13	47	26	29	15	181	99
NARA's Guidelines for Record Creation	82	45	32	18	63	35	4	2	181	100
OMB Bulletin 95-01	86	48	30	17	62	34	3	2	181	101
Z39.50 Standard	38	21	45	25	92	51	6	3	181	100
FIPS No. 192	41	22	30	17	106	59	4	2	181	100
PRA 1995, GILS Section	81	45	45	25	53	29	2	1	181	100
The World Wide Web	121	67	30	17	28	15	2	1	181	100

* Total does not equal 100% due to rounding.

Potential users contacted by the study team often did not know of, nor had they used, GILS. Indeed, a primary audience and potential user group, government document librarians participating in an early focus group, revealed very little knowledge of GILS. Further, this user group's interest is also limited, if assessed on the fact that only one person attended a focus group of documents librarians at the American Library Association Midyear Conference in February 1997. GPO staff, finding little awareness of GILS in their training of depository librarians, developed and are delivering training on GILS as part of the overall training on GPO Access. This targeted training effort should improve awareness among the documents librarian community.

4.6.3. FINDING: GILS Usage Is Limited

The study identified a generally low level of awareness of GILS. As part of the study, several tactics were used to gauge current use of GILS. One technique was the GILS Conference survey; another was the transaction log analysis—the data from the latter requiring caution in interpretation. The findings from the survey pointed out quite clearly that GILS Conference participants, who, after 2 years into the initiative, should be knowledgeable and aware of GILS (see Table 4-10), revealed that their actual use of GILS is very low.

The survey asked GILS Conference respondents to indicate their uses of GILS in a series of True/False statements. A large majority of respondents neither use GILS frequently nor do they refer people to GILS for finding information. A majority (54%) do not find useful information when using GILS. Responses to these questions—especially given the nature of the respondents (i.e., primarily Federal agency staff with some interest in GILS)—raise the question as to the usefulness of GILS for these respondents. Table 4-11 summarizes the responses to these statements.

Another perspective on GILS comes from transaction logs for searches and hits against Web servers. Most of the GILS implementations sit behind Web (i.e., HTTP) servers, and often are mounted as a database on a WAIS server. It is possible to capture search and retrieve transactions on both the WAIS and Web servers to obtain an indication of GILS use.

GPO compiles and publishes a summary of monthly GPO Access GILS Usage Statistics (available from <[gopher://gopher.cni.org:70/11/cniftp/pub/gils/form](http://gopher.cni.org:70/11/cniftp/pub/gils/form)>). Table 4-12 presents a summary of searches on its GILS site since April 1996. The numbers reflect searches against the GILS database, and GPO provides the following to indicate what these numbers include: "the database listed as 'GILS', represents searches performed when a user chose to search all agencies' records."

Table 4-11
Use of GILS

Use of GILS	True		False		Total	
	N	%	N	%	N	%
I find useful information when I use GILS	84	46%	97	54%	181	100%
I often find links to GILS on the Web	53	29%	128	71%	181	100%
I often refer people to GILS when providing information	48	27%	133	73%	181	100%
I search GILS several times per day in my everyday work	6	3%	175	97%	181	100%

With Table 4-12, one should note that the "Difference from Average Searches" (and similarly for retrievals) needs to be interpreted carefully because of the variance of minimum and maximum searches and retrievals. In addition, the average number of searches and retrievals reflects the strength of the extreme maximum and minimum outliers (i.e., April 1996 and December 1996, respectively) on the average. One can also look at these numbers for an indication of the trend of GILS usage on GPO Access.

GPO provides statistics not only for its GILS database, but for all agency GILS databases it hosts, and a similar table could be generated for each of those databases. Individual agencies, as well as GPO, can use such statistics to analyze access to and use of their GILS records.

The evaluation study also used log analysis procedures of HTTP server transaction files. Appendix E-4 summarizes this analysis, which was an exploratory procedure. The data reflect a two-week period of transactions on one agency's server. Since the agency's GILS database sits behind the HTTP server, the server statistics capture all HTTP transactions (i.e., transactions against all the resources, including the GILS database, that are accessible through the HTTP server). This configuration allowed the study team to estimate the amount of GILS usage as a percentage of total Web transactions on the agency's server. Table 4-13 summarizes the analysis and indicates that GILS activities accounted for less than 1% of all Web transactions.

The data in Table 4-13 should not be compared to the numbers for GPO GILS searches in Table 4-12, as the two sets were collected in and reflect entirely different contexts. GPO's data resulted from the context of searches against its WAIS server, while Table 4-13 reflects numbers of hits and accesses in the context of an agency's HTTP server.

Further, the numbers in Table 4-13 should be interpreted with caution. The agency's HTTP server provides access to a rich collection of documents and other online resources and services. One possible type of analysis would be to identify the "percentage" of resources on the HTTP server represented by the GILS records and then see if the percentage of GILS hits is commensurate with the percentage of resources represented by GILS. While it is possible to state from this log file data that GILS hits and accesses relative to overall server hits and accesses comprise a very small percentage, it would be unwise to conclude that the small percentage of GILS hits and accesses represents low utilization of GILS compared with utilization of other resources accessible via this HTTP server. Longitudinal data over months could, however, reflect whether use of GILS resources on the HTTP server is increasing, decreasing, or remaining steady.

As noted above, from the transaction log analysis carried out during a two week test period, hits to the one agency's GILS constituted less than 1% of all hits to that agency's HTTP server. The DTIC GILS locator page on their Web indicated some 34,000 hits during October, 1996. DTIC officials estimated that hits on GILS are also less than 1% of all DTIC Web server hits (the URL for DoD web statistics is:

<www.dtic.mil/dusage/>). These individual agency estimates of GILS usage cannot be generalized to overall GILS use, but they do provide one indication as to its use.

The study could not identify any reports of GILS use by agencies except for the published GPO statistics and some basic HTTP log analysis statistics. This suggests that as part of agencies' responsibilities in a refocused GILS, they will need to report as one performance measure an estimate of GILS use (in whatever ways they measure it).

One anecdote regarding use is especially instructive. The librarian at a Federal agency responsible for inputting GILS records told the investigators that she rarely if ever used the tool for

identifying and accessing government information—especially since she knew the type of records that were being input! Interestingly, most agency participants in the study, including this librarian, agreed with the list of problems identified in the user assessment. They favored letting people first become familiar with it, completing an assessment (such as that reported here), and then deciding how to improve GILS.

The findings from the study also indicate that without a substantial investment of time and resources in education and promotion, the investigators question the extent to which people will become knowledgeable about GILS to say nothing of them becoming familiar with it and using it regularly.

Table 4-12
GPO GILS Usage Statistics

Month-Year	Searches	Difference From Average Searches	Retrievals	Difference from Average Retrievals
April 96	20,453	+41 %	22,154	+55 %
May 96	13,975	-3 %	20,174	+42 %
June 96	13,878	-4 %	14,030	-2 %
July 96	13,147	-9 %	12,223	-14 %
August 96	12,773	-12 %	11,860	-17 %
September 96	14,213	-2 %	12,600*	-12 %
October 96	17,420	+21 %	15,674	+10 %
November 96	13,099	-9 %	11,433	-20 %
December 96	11,690	-19 %	11,834	-17 %
January 97	13,840	-4 %	12,436	-13 %
February 97	11,988	-17 %	10,971	-23 %
March 97	16,995	+18 %	15,658	+10 %
TOTAL	173,471		171,047	
Minimum	11,690		10,971	
Maximum	20,453		22,154	
Average	14,455		14,253	

*Source read "1,260"; assumed correction as shown.

Note: GPO provides the following definitions of search and retrieval:

"A search is counted each time a particular database is queried. A retrieval represents a file actually being transferred onto a local machine, as opposed to frequently reported "hits" statistics, which represent each mouse click or change of a Web page."

BEST COPY AVAILABLE

Table 4-13
GILS Hits* and Accesses as Percentage of Agency HTTP Server

Week	Hits on HTTP Server	Hits on GILS Database	% of GILS Hits on HTTP Server	Accesses on HTTP Server	Accesses on GILS Database	% of GILS Accesses on HTTP Server
Week 1 (2/2/97 - 2/8/97)	1,688,596	3,844	0.22%	569,326	2,977	0.52%
Week 2 (2/9/97 - 2/15/97)	1,496,127	4,824	0.32%	564,776	3,451	0.61%

* A *hit* is any file from a web site that a user downloads. A hit can be a text document, image, movie, or a sound file. If a user downloads a page with 6 images on it, then that user "hit" the web site seven times (6 images + 1 text page).

An *access*, or sometimes called a page hit, is an entire page downloaded by a user regardless of the number of images, sounds, or movies. If a user downloads a web page that has 6 images on it, then that user just accessed one page of the web site.

4.6.4. RECOMMENDATION: Develop and Formalize GILS Promotion, Education, and Training Strategies

The original GILS effort as outlined in OMB Bulletin 95-01 charged NARA with responsibility for training—primarily in the area of record creation and maintenance. NARA developed guidelines for record creation and offered a number of training sessions. The study found the need for *more* than training on creating records. In fact, the investigators found a need for an education and awareness *program* directed at agency management and agency implementors that would describe and explain how GILS could assist them and what benefits would accrue from participating in GILS. To do this, however, would have required a clearer articulation of the purposes and goals of GILS, which can be achieved in a refocused GILS initiative.

The investigators were unable to identify a *government-wide* marketing plan or program for the development and implementation of GILS. The investigators did, however, identify some informal efforts within some agencies

First, policy leaders should understand that education and marketing efforts have two very different target audiences: Federal agency staff and users/potential users of GILS. Findings from the study show clearly that neither of these two groups understand the purpose, importance, and potential

benefits of GILS. The educational, training, and promotion objectives for each of these groups need to be customized both in content and in delivery. Further, the category of Federal agency staff includes different subgroups such as agency management and actual GILS implementors (and potential GILS users), and education and marketing for these subgroups may have different emphases.

Second, these promotional efforts cannot be planned and implemented until GILS policymakers articulate a clear, achievable purpose, define specific objectives, and agree to implementation procedures that would constitute a refocused GILS. A major problem that developed during the early period of Federal GILS implementation was that various GILS spokespersons oftentimes provided differing visions and purposes for GILS efforts.

In addition, clear lines of responsibility for these efforts need to be established. There are a number of options related to developing a coordinated effort. Assuming the CIO Council takes on (or is charged with) overall GILS development, a Council GILS Committee or interagency task group should have responsibility for a *program* of planning, implementing, and evaluating a promotional effort. The term *program* is used to stress the importance of this effort being ongoing and credible.

While there are numerous ways to cast a refocused GILS marketing, education, and training effort, such efforts require careful attention. A formal

mechanism should be established to plan and implement a marketing, education, and training effort as described above. Specific objectives should be developed for specific target audiences:

- Establish a procedure for an "official" spokesperson for the refocused GILS efforts with "official" oversight as to the content of news releases, brochures, etc.
- Document and demonstrate to government officials "best practices" implementation of GILS.
- Demonstrate to government officials specific benefits that will result from a refocused GILS implementation.
- Ensure that the public and more specifically, targeted user communities of GILS have accurate expectations of its products and services.
- Develop brochures promoting the use and importance of GILS.
- Encourage agencies to mount prominently on each GILS site standardized statements that clearly articulate the GILS mission, operability, limitations, and instructions for use.

These objectives are illustrative only. Discussion among policy leaders will need to occur as to how best to plan and implement a *program* of promotion for a refocused GILS effort. Regardless of the approach taken, these efforts should stress the refocused GILS as a *government-wide* tool and *not* simply an agency-based tool.

4.7. TRANSITIONING TO THE NEXT STAGE OF GILS DEPLOYMENT

In *Access America: Reengineering Through Information Technology*, Vice President Gore states "Information Technology (IT) was and is the great enabler for reinvention. It allows us to rethink, in fundamental ways, how people work and how we serve customers" (National Performance Review and the Government Information Technology Services Board, 1997, p. 1). The original vision of GILS, while appropriate at the time, is in need of reengineering. The degree to which this

reengineering process is tied to clear purposes and objectives for GILS will dictate the success of the refocused GILS effort.

This chapter reported a number of findings concerning the "success" of GILS. These findings indicate that the GILS vision as outlined in OMB Bulletin 95-01 has not been reached despite some individual agency successes. Beyond that vision, however, the study found a desire to articulate a refocused GILS vision, more in keeping with the networked environment in which GILS is deployed. The refocused GILS builds on the basic architecture of decentralized agency-based databases of structured metadata records accessible via Z39.50. The refocused GILS is clearly an evolutionary step in GILS development.

The findings and recommendations offered in this chapter cover a very broad range of topics and issues. These findings and recommendations describe an initial GILS implementation effort that has had mixed results as of this writing. But these mixed results provide a richness in lessons learned that can guide a refocused GILS effort. Indeed, one might suggest that an initial period such as that during 1995–1996 is inevitable when implementing a complex and multifaceted program such as GILS. The findings and recommendations, while important, may be less important than the resolve to learn from them and develop a clear path to the next stage of GILS deployment.

Clearly there are issues yet to be resolved in refocusing GILS. Yet those issues are the catalyst to move GILS forward and continue to learn and improve subsequent efforts. There has been a significant amount of knowledge gained during this GILS implementation effort. This report, and more specifically this chapter, documents that knowledge and offers recommendations to *build* on this knowledge. The investigators believe that the notion of a U.S. GILS is still very powerful, one that if refocused appropriately has the potential to make significant improvements in accessing government information and managing government information resources.

Chapter 5 will discuss the nature of this transition period, and offer some possible strategies for

sequencing recommendations to deal with those that are most important. During this transition period, however, a carefully developed plan with individuals or agencies clearly responsible for project management of the transition is needed. The transition team that manages the transition will need to investigate a number of additional research questions that Chapter 5 outlines. Until some of these research questions are addressed, resolution of key GILS issues will be problematic.

The knowledge gained from the initial GILS implementation is significant and useful. This

knowledge, as outlined in this report, should inform the refocused GILS initiative. The investigators expect discussion and debate about the findings and recommendations offered in this chapter. Clearly, not everyone will agree with all the recommendations offered. More important, however, is that there is a *clear and agreed upon* vision of the refocused GILS effort and that careful planning guide the transition to the next stage of GILS. This chapter informs that transition process, and Chapter 5 offers some guidelines and research questions to be considered in moving to a refocused GILS.

Chapter 5

Priorities,

Next Steps, and

Future Research

5.0. INTRODUCTION

The findings of this evaluation study indicate that GILS has not reached its full potential as a *government-wide* information locator service. A number of agencies, however, have put their "agency GILS" to work in interesting ways including assisting in public access to government information and broader information management efforts. In addition, the basic concept of GILS as a set of decentralized, agency-based locators containing structured metadata records and accessible via Z39.50 remains a valid architecture to support networked information discovery and retrieval. Yet, what exists in Spring 1997 is not a government-wide information locator service but a set of diverse agency implementations that vary in coverage and scope. The past two years of implementation experience highlight important issues—at both policy and implementation levels. Without this experience, neither agencies nor users would be able to articulate the issues identified in this study.

Networked-based locator services such as GILS should be seen as innovative approaches for providing access to government information. Precisely because of its innovative character, the U.S. Federal GILS initiative has identified and clarified some basic research issues for networked information discovery and retrieval. In fact, aspects of GILS such as the capture/use of metadata and distributed search and retrieval tools are essentially research issues for which scalable and operational solutions have yet to be fully developed. The recently released report on the Canadian GILS pilot project (see Appendix I) parallels many of the findings and recommendations reported in this study of U.S. Federal GILS implementation. A comparison of the two reports indicates that networked-based locator services share systemic problems and common issues. In the view of the investigators of this study, the common threads that run through both reports point out that the architecture of GILS provides a valid approach to a networked-based locator service, yet carrying out that architecture in actual implementation makes visible important research questions.

Two years of agency GILS implementation experience, however, provide both the implementors

and policymakers with a firm basis for determining the future shape of a government information locator service. One consequence of this study has been documenting the range of issues, problems, and success factors that are only visible now because of the actual implementation experience. The findings and associated recommendations discussed in Chapter 4 contribute to the discussions among policymakers and implementors in deciding what next steps need to be taken and how those steps should be sequenced. The recommendations in Chapter 4 range in detail and priority from, for example, government-wide information policy integration to specifics about the data elements in GILS records. This chapter provides a summary of the recommendations and places them in a framework for action.

One of the first actions resulting from the completion of this report will be the responsibility of the GILS evaluation advisory group. The GILS Board charged a group of agency representatives with planning the evaluation study. Members of that committee have served as an advisory group to the investigators during the study. When this final report is submitted to the COTR and distributed to the advisory group, that group will have the responsibility to meet, review and respond to the findings and recommendations in this report, and decide on specific next steps to move the GILS initiative forward. This report will serve as a point of departure for discussions and agreements among the advisory group, and the advisory group's actions will guide the next phase and the overall success of the U.S. Federal GILS initiative. The recommendations and framework for action reported below provide the advisory group with a beginning point for their deliberations.

In the final assessment, there is much left to learn about networked-based locator services. Throughout the report, the investigators described many issues and problems that are beyond the scope of this report to resolve, and often the recommendations identify areas in need of further research. This final chapter enumerates a series of research areas that require attention if the refocused GILS initiative is to succeed.

5.1. CHARACTERIZING THE REFOCUSED GILS: RECOMMENDATIONS

This report has documented major issues facing the U.S. Federal GILS initiative. If a true *government-wide* information locator service is to evolve from the agency implementations, the investigators have concluded that major revisions in the approach to GILS are necessary. These revisions are at policy, administrative, oversight, and implementation levels. The investigators recommend that such revisions should occur based on a refocusing of the purpose, goals, and scope of the current approach to GILS.

A refocused GILS initiative builds on the success of selected agency GILS implementations, but with a clear demarcation from the current approach. The investigators concluded that given the current confusion over what GILS means, it is essential that policymakers and implementors clearly differentiate the refocused GILS from the effort guided by the original OMB 95-01. The new initiative acknowledges the value of many aspects of the original GILS concept, yet policy for a refocused GILS provides a clear line of demarcation between the early GILS implementation period (i.e., 1995-1996) and a refocused GILS. One approach to distinguishing the refocusing of GILS is through a change in the name to reflect, for example, a "second release" of the U.S. Federal GILS service. An important aspect of such a demarcation with the early implementation period is to acknowledge the lessons learned from that experience. To give some indication of what a refocused GILS could include, the investigators offer the following major recommendations.

5.1.1. The Refocused GILS Initiative Clearly Articulates the Purposes and Utility of a Government Information Locator Service

The current GILS is different things to different people and has led to inconsistent implementations and a wide range of expectations of GILS. A refocused GILS must clearly articulate the function of a *government-wide* information locator, its scope

of coverage, what people can legitimately expect it to provide, and the benefits it offers.

The investigators recommend that GILS be refocused and aligned with the following vision:

An easy-to-use and coherent government-wide information search service available from one or more service points that enables users to discover, locate, select, and access publicly available government information resources (e.g., agency information systems, specific information dissemination products, and existing locators to those products) through standardized metadata that describe those resources and provide direct links to the described resource (e.g., full-text documents, other online services).

The purpose of the refocused GILS is to enable users to discover what government information exists and provide users with direct access to that information. The revised purpose does not include records management. Any additional functions proposed for GILS that extend this initial purpose must be tested and demonstrated prior to raising expectations and to determine whether or not GILS can achieve or support such purposes and functions (e.g. EFOIA).

To support networked information discovery and retrieval, the GILS records (i.e., metadata) will be crucial. Web-based searching or browsing as currently offered by agency Web sites or through implementations such as the White House or Thomas Web sites do not give users a government-wide view nor provide government-wide discovery and access. GILS records are a necessary lynchpin to solve the networked information discovery and access problem.

The scope of a refocused GILS should a subset of all government information resources, namely those that are in digital form. Due in part to the ease of Web publishing and the ease of interfacing with existing online databases and services through Web scripting, the amount of network-accessible government information will continue to increase. This scope of coverage is *realistic* rather than simply *reasonable*. First, users would know what they could expect to find if the scope of the refocused GILS is "metadata

records that describe the publicly accessible electronic resources of the government and provide linkages or access to those resources." A refocused and more limited scope would provide guidance to agencies in their development of GILS implementations, especially by clearly specifying the agency resources GILS records should describe. Also, this scope accounts for the increased expectation of users who want to obtain the *actual* information rather than just a description of it.

The GILS that results from a systematic refocusing can clarify to agencies and users what the government information locator service is, how it works, what is covered, and what users can expect from it.

5.1.2. The Refocused GILS Initiative Provides Clear Lines of Authority and Oversight

An essential feature of U. S. Federal GILS is its decentralized approach—at the agency level—for providing locator services to agency information resources. To date, this decentralized implementation responsibility has not been balanced by integrated or coordinated management and administration. The refocused GILS initiative—through policy directive—identifies an appropriate organizational unit that has the responsibility, authority, and accountability for providing government-wide coordination and administration of GILS activities. The refocused GILS initiative, however, safeguards the decentralized character of agency-based locators, where those people closest to the resources are responsible for identifying them, assisting in the creation and maintenance of GILS records, and providing public access to them.

The investigators recommend that two organizational units be charged with separate mandates for the next stage of GILS development and deployment. First, the GILS Board should provide the forum—through various task forces—for determining the revised focus of GILS. The result of discussions by the task force and the GILS Board should be a set of policy recommendations submitted to the Office of Management and Budget as a basis for its revisions to OMB Bulletin 95-01 that expires at the end of 1997. The forum provided by the GILS Board should

include representatives of agencies as well as non-Federal stakeholders in GILS such as citizens, librarians, researchers, and public interest groups.

Second, the CIO Council is an appropriate interagency body that can provide ongoing coordination and administration of GILS. The government-wide character of a government information locator service combined with the Council's mandate makes it a suitable locus for these responsibilities. One important function of the Council will be setting appropriate, realistic, and measurable objectives for agency GILS implementations. OMB Policy directives and goals for the refocused GILS need to be translated into actionable steps that agencies can take. Measurable, specific, and realistic objectives can guide agency actions. In addition, the Council should determine reporting schedules, receive agency GILS progress and implementation reports, and provide information to OMB and the GILS Board on the status of GILS activities. The CIO Council should have representation on the GILS Board to enable communication between the two organizations.

5.1.3. The Refocused GILS Initiative Demonstrates Effectiveness and Benefits Through a Pilot Program

The refocused GILS initiative recognizes that networked information discovery and retrieval (NIDR) is a new and as yet relatively unexplored terrain. Many of the current technical implementation issues are part of the larger research area dealing with NIDR. Systems for the organization and access to information—government or otherwise—have developed over the years, often through experimentation and lessons learned from implementation experience. The networked environment, which is the context for all information handling at the end of the 20th Century, adds new layers of complexity to traditional approaches of information organization and access. The refocused GILS initiative participates fully in utilizing emerging information technologies to improve access to government information. Further, the refocused GILS initiative should acknowledge the immaturity of NIDR by establishing a ongoing pilot program to

identify problems and issues in both policy and implementation arenas.

The investigators recommend that once stakeholders reach consensus on the character and specifics of a refocused GILS, the CIO Council will establish a GILS pilot program. A GILS pilot and demonstration program offers many benefits to implementing agencies and users. Agencies gain the benefit of tested technology, procedures, and best practices. Pilot implementations can demonstrate tangible benefits to those agencies that need convincing that GILS is worth doing, and doing well. Users can experience the utility of a government-wide search and retrieval service. More importantly, users can provide critical input at the design and development phases of the next generation of GILS implementations through a pilot program to ensure that the resulting information locator service meets the requirements of various user communities that need access to government information.

Working in parallel, OMB, the GILS Board, and the CIO Council should establish policies, goals, and objectives for the refocused GILS. Specific objectives will provide a standard against which the implementations in the GILS pilot program can be measured. Thus, the pilot program serves as a check on the ability of GILS implementations to achieve the objectives, goals, and policies of the refocused GILS initiative.

A pilot program can serve as a testbed for experimental implementations of any additional functionality that policymakers and agency implementors determine *reasonable* for the refocused GILS. The important point is that before policymakers or implementors raise expectations of functionality to be supported by a discovery and access service such as GILS, experimentation occurs to test and demonstrate how *realistically* GILS supports additional functionality. A pilot program also can be effective in determining the prospects of new and emerging technologies, such as the proposed Advanced Search Facility (e.g., single-point search and retrieval) or the current push technology (e.g., for announcing updates of GILS services to cognizant communities).

Most importantly, a pilot program serves as the focal point for resolving many of the technical and procedural issues identified in Chapter 4 including the appropriate data elements in a GILS record to support discovery and access, the most efficient procedures for capture of metadata (manual or automatic), and the presentation of GILS records to users. To ensure that the results of a pilot program serve all agency implementors, the program should include representative implementors from large and small agencies. In cases where smaller agencies face severe resource constraints for participation, incentives and funding may be necessary.

5.1.4. The Refocused GILS Initiative Includes a Program of Government-Wide Education and Promotion

The refocused GILS initiative rejects the *If we build it, they will come* perspective. Policymakers and implementors should promote GILS as the first point of contact for users looking for government information. A clearly articulated purpose for a refocused GILS, aligned with demonstrable utility of it through a GILS pilot program, will be a basis for developing a program of promotion and education.

The investigators recommend that if GILS is worth doing, it is worth promoting! In the networked environment, there are many competitors for the attention of information seekers. GILS is a service that can compete, since it offers the general public with a government-wide search and retrieval service not offered by other online services. The Government has no *special advantage* in the marketplace with GILS; GILS is a service that offers a product (freely available) to other networked services providers to create value-added products of their own. GILS, however, is a service for which the Government has special responsibilities, since it will be the point at which citizens and the Government intersect for information access and dissemination.

A government-wide program of education and promotion also includes a focus on the agencies themselves. Agencies need to better understand what the refocused GILS can offer them, and a program of

education directed at the agencies can build intra-agency support for the next stage of GILS development. A GILS pilot program will demonstrate tangible benefits as well as provide tested practices and procedures to the agencies. That effort may result in improved agency management buy-in, which may in turn result in sufficient and dedicated funds for agency GILS activities.

5.1.5. The Refocused GILS Initiative Emphasizes Continuous Improvement Through Ongoing Evaluation

The refocused GILS initiative acknowledges the evolutionary character of the networked environment, the changing needs, behaviors, and expectations of users, and the need for GILS to evolve to address user requirements and technology changes. GILS should be committed to the goal of continuous improvement with a resulting service and product that is responsive to its users. Its focus on users requires ongoing evaluation and assessment by the users for which it was developed.

The investigators recommend that ongoing evaluation be a essential component of the refocused GILS. As reported from this study, GILS is a complex, networked service that can be assessed along multiple dimensions and from multiple perspectives. In addition, this study demonstrated tools and procedures for assessing various dimensions of GILS. Ongoing evaluation of GILS must incorporate a user-based approach since the users—internal agency users or external citizen users—are the final arbiters of the success of GILS. The refocused GILS must also identify specific and measurable objectives against which it can be assessed. Therefore, evaluation programs and procedures need to be incorporated during the early discussions about the refocused purpose, scope, functions, and objectives of GILS. GILS policymakers and implementors need to understand the objectives and criteria that constitute a successful GILS, and they must understand the need for and be able to collect appropriate information to conduct useful assessments.

5.2. PRIORITIES FOR ACTION TOWARDS THE REFOCUSED GILS

The characterization of the refocused GILS and associated major recommendations discussed above do not address the details of the issues documented in this report. Rather, they suggest an initial set of actions for government and agency policymakers and implementors to move toward specific solutions and to encourage the success of the refocused GILS across the Federal government. The investigators imagine a number of possible scenarios for the evolution into the next stage of GILS deployment, but all include the identification and prioritization of actions by policymakers and implementors. This section briefly outlines one ordering of priorities based on the findings and recommendations reported in Chapter 4. In a certain sense, the following ordering reflects a commonsense approach in that certain actions and decisions occur logically before to others.

The highest priority for the refocused GILS initiative is to gain consensus on the purposes, goals, and scope of GILS. The investigators have concluded that a major flaw with the current GILS effort is a lack of clearly understood purpose and utility. Chapter 4 noted that the successful agency GILS implementations occurred where agencies determined what GILS would be for them regardless of whether it aligned exactly with the prescriptions of OMB Bulletin 95-01. By defining a clear purpose for GILS, these agencies also identified its utility and recognized the benefits from their GILS implementations. Such agency-centric approaches, however, are unlikely to lead to a coherent *government-wide* information locator service. The refocused GILS must be directed by policy, but that policy needs to be built upon the consensus of individuals representing the affected agencies *and* the public that GILS serves. The first and most important step is to articulate the purpose of the refocused GILS, what it is, how it will work, and the potential benefits that will accrue to agencies and the public.

The second priority is for GILS policy to state clearly who has authority, who is responsible, and where accountability will rest for the refocused GILS as a *government-wide* initiative. This needs

to include explicit statements concerning what such authority and responsibility entails. OMB Bulletin 95-01 named several agencies with various responsibilities for GILS, but except for the GILS Board, no organizational unit had government-wide authority, responsibility, or accountability. By devolving all GILS activities to the agencies without overall coordinating counterbalance, the result was very uneven implementations or no implementations at all. Part of this coordination and administration responsibility is to provide a necessary forum where agency implementors and others can work out specific implementation issues, requirements, and strategies.

The third priority is to develop policy goals for the refocused GILS and translate them into specific, realistic, and measurable objectives. OMB develops policy in consultation and with advice from agencies, the GILS Board, and others. OMB voices the information policy goals for the Federal government. OMB, however, does not have the responsibility for micromanaging the agencies, and the translation of policy goals to specific objectives must be carried out by others. One level at which this can happen is the organizational unit responsible for government-wide coordination and administration of the refocused GILS. At the agency level, appropriate objectives for GILS will also be developed. Without the intervening government-wide coordination level, however, agencies may take too much latitude in interpreting OMB goals and translating them—for themselves—into objectives that do not support the *government-wide* character of refocused GILS.

The fourth and final priority for initiating the refocused GILS effort is to establish a GILS pilot program. The organizational unit responsible for government-wide coordination could be charged with overseeing and administering a pilot program. To maintain a government-wide perspective for the refocused GILS, representative agencies of all sizes and missions should be included in a pilot program. A pilot program does not have to result in technologies, procedures, and practices that are a “one size fits all,” and the variance of agency missions and resources must be reflected in the participants in the pilot.

The investigators recommend that these four priorities are critical first steps to move to the next stage of GILS evolution. The investigators also recommend that the GILS Board, with advice from the CIO Council and OMB, establish a GILS Transition Task Force to address these priorities.

5.3. THE POLICY CONTEXT FOR THE REFOCUSED GILS

As discussed earlier in this report, the Federal policy context for information management has been dynamic and challenging in recent years to say the least. The specific context for GILS, as described in Chapter 2, mirrored this phenomenon. There is some sense that the efforts of the Information Infrastructure Task Force (IITF) on GILS, OMB Bulletin 95-01, NARA's *Guidelines for the Preparation of GILS Core Entries*, and other efforts to implement GILS have been a grand experiment from which a significant amount of knowledge has been gained to improve the existing GILS context and refocus the service to accommodate evolving expectations.

There is wide agreement that strong and visible support from the OMB-OIRA, the CIO Council, the GILS Board, the Government Information Technology Services Board (GITSB), and the office of the Vice President as part of the National Performance Review are critical to the successful evolution of GILS (however it might be recast or reinvented). There is equal agreement that such support does not currently exist. Also important is the need for a better understanding of the roles and responsibilities for *policy leadership* as well as *implementation leadership*.

The study finds that OMB Bulletin 95-01 was a good first effort to outline a policy context for the development of GILS. Some issues that will require attention in a forthcoming revision to the Bulletin include:

- Clarifying purpose and objectives of GILS (e.g., relationship with EFOIA, if any)
- Divesting records management responsibilities and activities from GILS

- Clarifying Federal leadership for a range of GILS activities
- Recognizing the extent to which agencies can take on GILS responsibilities in a time of budget reductions and increased demands on productivity
- Indicating realistic and tangible benefits that can accrue from GILS
- Integrating GILS into a broader context of agency information systems (including Web sites), IRM, and general information management missions
- Providing regular oversight and enforcement of GILS policies
- Promoting the development of search and retrieval mechanisms and processes that integrate and coordinate agency components of GILS into a *government-wide GILS*.

These areas for policy revision are illustrative only. They do, however provide a flavor of the range and content that will need to be addressed in a revised OMB Bulletin on GILS. The investigators suggest other concerns (see Chapter 4) that will also require attention in a revised Bulletin.

5.4. FURTHER RESEARCH AND EVALUATION EFFORTS

The study collected a significant amount of data that describe many GILS-related activities and products. This evaluation also identified areas of policy and implementation needing additional research. Such research should be conducted to improve understanding of how future efforts broadly related to organization, discovery, and access to government information can be improved. GILS policymakers and implementors need to recognize, however, that networked information discovery and retrieval (NIDR) is basically still an evolving research area. This report noted that scalable and operational solutions to issues related to NIDR have yet to emerge. A GILS pilot program offers a valuable opportunity to conduct research on issues specific to U.S. Federal GILS implementation. Some specific research areas in need of additional attention include:

- **Use, Presentation, and Content of Metadata Records:** Findings from this study and knowledge of other networked information discovery and retrieval activities suggest that metadata records may be a critical component in utilizing networked resources. For the refocused GILS initiative, a formal assessment should be done of the appropriateness of the metadata, the use of each of the data elements, and the extent to which they can support the goals of GILS. This research should include the use of metadata by human as well as machine processes (e.g., software agents). Alternative approaches for visually presenting the GILS metadata should be developed and tested.
- **Linkages between GILS and Web:** This study has identified a number of approaches by which GILS or metadata information can be integrated and linked to agency Web sites. Research into which types of approaches are most effective (from a user as well as developer perspective) should be initiated. Criteria can be developed against which the various approaches can be assessed.
- **Extent of Coverage of GILS:** Assessing the existing coverage of agency information resources by GILS was beyond the scope of the current study. Such assessment is necessary, however, to indicate agency compliance with OMB Bulletin 95-01. Research on this topic would examine the appropriate units of analysis for GILS records (i.e., specifying the aggregation/granularity of objects suitable for description), as well as identifying the universe of agency resources that should be described. To assess extent of coverage, however, would require knowledge of all agency information resources. This research would assist in answering the question: how many GILS records are sufficient and appropriate to provide coverage of all agency information resources.
- **Networked Services and User-Performance Variables:** Another aspect

of networked information discovery and retrieval is isolating user-performance variables relative to the environment (the Web) versus the functionality of system design. In a number of the assessments conducted (e.g., scripted online user assessment), it was difficult to determine whether, for example, poor response time is due to difficulties in Internet routing, technical design of the agency GILS, poor server response at the agency, or other variables.

- **Cost Benefit Studies:** Some agencies that participated in this study clearly believed that the current GILS initiative was well worth the effort in light of the costs and other various resources committed to the effort. Others were adamant that no benefits occurred regardless of the costs involved. Additional study into why there are such vast differences in perceptions would be very instructive. Indeed, simply being able to identify *specific* costs directly associated with the GILS effort on an agency by agency basis would be useful. To a large degree, the investigators have to take at face value participants' views of costs and benefits with little supporting "evidence."
- **Performance Criteria and Indicators:** Neither OMB Bulletin 95-01 nor agency implementors detailed performance criteria and indicators, thus it is very difficult to determine, *post-hoc*, the degree to which a particular GILS effort can be described as a "success." The next phase of GILS should include a research initiative to determine a number of performance measures that can be used to gauge the success of the effort. Such performance measures require development, testing, and validity assessment (McClure & Lopata, 1996; National Academy of Public Administration, 1996). Further, this research could address how GILS links to the *Government Performance and Results Act* (GPRA). The Act requires that all agency programs have performance measures. GILS may be seen as an

- enabling tool for assessing program accomplishment, and it also needs to be assessed as a program itself.
- **Government-Wide Search and Retrieval:** Effective searching across agency GILS (and other databases) or otherwise massive amounts of data requires additional research. In part, the failure of GILS to provide an effective cross-agency search facility is limited by the availability of *effective* search and retrieval tools. Support for distributed search and retrieval technologies such as the Advanced Search Facility (ASF) or some ASF-like effort is essential. Further, recent applications by GPO and FedWorld for cross-agency searching need to be evaluated to measure their utility for government-wide searching. User-based assessments on search and retrieval tools are required.
- **Agency Staff Responsibilities, Accountabilities, and Authorities:** The study identified a wide range of agency staff and offices that ultimately had responsibility for "GILS-related activities" during the 1995-1996 effort. However, a systematic identification of who, specifically, had what types of responsibilities (i.e., management, records development, technical design, etc.), other responsibilities of these individuals, and some background information as to their education, experience, knowledge, and degree to which they had training, would be very useful. Such information could help explain, possibly, the significant discrepancies in agencies' perspectives toward the GILS effort.
- **Comparative Studies of Other Government Information Locator Service Implementations:** Numerous state and international GILS implementations are occurring (e.g., Washington State and Canada). Although the U.S. Federal implementation of GILS preceded others' efforts, analysis of policy and procedures of these non-Federal

- implementations could provide practices and procedures to improve U.S. Federal activities.
- **Government-Wide Records Management System:** This study recommended that NARA be tasked with the responsibility for developing a records management system to better identify, schedule, and ultimately preserve appropriate government information resources, especially resources in electronic or digital form. This effort will require a research component *prior to* any system design and testing. The component will need to clearly identify system requirements, determine the agency uses and applications of such a system, and describe existing techniques for management of electronic records.
- **Policy Review and Analysis of "Locator" Systems:** The literature and policy review provided in Chapter 2 identifies a range of ambiguous, contradictory, and confusing policy language related to government-wide locator systems. Policy research and analysis is needed to develop one coherent statement that organizes policy language from these various instruments. The results of such research can be included in a future revision of OMB Bulletin 95-01.

This list of additional areas requiring research is not intended to be comprehensive. Rather it is illustrative of key topics from which additional knowledge would be extremely helpful in supporting possible future GILS activities.

Another thrust of this study was to design, develop, and test assessment techniques. The intent was to provide policymakers and agency officials with *tools* by which *they* could deploy a range of assessment techniques and comply with policy such as GPRA. To date there has been little consideration (at least as identified in this study) about agency-based performance assessment and the development of performance indicators for GILS efforts.

The various instruments developed for this project should be seen as first efforts. Additional research related to these evaluation tools is both necessary and appropriate. Some possible areas for additional effort, for example, include:

- **Log Analysis:** Appendix E-4 provides detailed explanations, techniques, and findings regarding the log analysis done for two weeks of HTTP transaction logs from EPA. There are numerous avenues of additional research in this area, some of which are outlined in the appendix. These techniques offer agencies and policymakers an important tool for monitoring and refining Web-based services. While some agencies do make use of basic log analysis techniques, most have yet to explore the techniques described or to develop the techniques proposed in this study.
- **User-Based Assessment Techniques:** The study found that the scripted online user assessment approach is a powerful tool for obtaining users' assessments of GILS information resources and services. Techniques developed here can be modified for use in individual agencies. Additional research should be undertaken on how to simplify the technique and how to better relate the scripting process to specific assessment criteria and performance indicators. In addition, video-taping users and asking them to "think aloud" as they use a particular networked service appears to have great potential as an assessment technique (Eschenfelder, et al., 1997). See Appendix C-5 for a description of the method used in this study and Appendix E-3 for the summary of results. Included in the latter appendix is a list of suggested questions and procedures to improve the method.
- **Metadata Record Content Analysis:** The study included a task that identified all GILS records and then analyzed the content of a sample of these records. This was an important first step (see Appendix C-4). An important next step will be to

refine the criteria and procedures for assessing the quality of metadata. Especially important is obtaining assessments of these metadata records from users. A scripted assessment technique, similar to that used to assess the agency GILS implementations could be developed for assessing the records. This research can identify primary, secondary, and tertiary metadata elements that support the purposes of GILS. Appendix E-2 identifies a series of questions that could guide research to improve this method.

Additional research on evaluation methods and data collection tools in the provision and management of networked information services is essential (McClure & Lopata, 1996). Overall, there has been little evaluation research in the area of government electronic networked services (Wyman, Beachboard & McClure, 1997).

An important benefit from this study is the development, testing, refining, and documentation of research techniques and evaluation tools. Initiatives related to GILS and its evolution should continue the development, testing, and use of assessment tools and methods discussed in this report. Indeed, these tools should help agencies better comply with GPRA for GILS-related activities and programs.

5.5. REENGINEERING THE GILS EFFORT

Policymakers must carefully determine the best approach to take for future GILS or GILS-like efforts. As reported in this study, there are serious issues and problems that currently limit the overall usefulness of GILS. The investigators believe that the original vision of GILS was not a clear one. GILS evolved into being different things for different people and agencies—evolving, with some notable exceptions, into an effort with little user or implementor support, limited usefulness, and with confounded purposes.

Having said this, however, one should not overlook the fact that a number of agencies developed

working and successful agency GILS, yet they defined the GILS vision within the context of their agency. These efforts are significant and should be recognized and applauded. Indeed, the knowledge gained by GILS implementations at DTIC, Treasury, EPA, GPO, and Interior, for example, is critical for the next stage of GILS development.

Despite these successes at a few agencies, the diagnosis is that GILS suffers from multiple-personalities disorder and schizophrenia. Despite its condition, GILS has not received adequate treatment from its "doctors" as they each have different views about the needs and appropriate treatments for the patient. Many individuals and agencies have given up on developing *any* treatment for it and have gone on to other more pressing problems. But such does not have to be the case for future efforts.

This study recommends that the existing GILS as developed during 1995–1996 be considered as Phase I. The lessons learned from this experience can contribute significantly to future efforts to develop a discovery and access service for government information. But GILS, as currently constituted and currently implemented, must be refocused and reengineered if it is to be a success.

A refocused GILS initiative based on the recommendations offered in this chapter is feasible and doable *IF* there is administrative coordination and commitment to completing such an effort and *IF* there is agreement as to the specific nature of the effort. The following vision for a refocused GILS could be a basis for such agreement:

An easy-to-use and coherent government-wide information search service available from one or more service points that enables users to discover, locate, select, and access publicly available government information resources (e.g., agency information systems, specific information dissemination products, and existing locators to those products) through standardized metadata that describe those resources and provide direct links to the described resource (e.g., full-text documents, other online services).

The next phase of GILS development will build on the basic architecture of decentralized, agency-based locators, standardized metadata records, use of Z39.50, and will draw as well upon Web and other technologies and developments in the arena of NIDR. The refocused GILS would provide *government-wide* search and retrieval capability and it would provide *direct* links to full-text information when available. It would provide online access to *information* and not just metadata, and the metadata records could be transparent to users except to provide them with characterization of resources that might be relevant to them. At the administrative level, the refocused GILS provides a balance between decentralized, agency-level GILS activities and government-wide oversight and coordination to result in a coherent and usable government-wide information locator service. This is only a broad brush at what that vision should be, but it offers a direction for the refocused GILS efforts. The investigators believe that the GILS experience can provide a significant number of lessons and information for moving forward with and improving GILS.

The vision of a tool that allows users to search for, discover, and obtain government information across all agencies in full-text via the network is an important vision to maintain—regardless of the future of GILS. While there is likely to be controversy and debate on how best to reach that vision, efforts should continue to make that vision a reality. Individual agencies cannot reach this vision on their own, however. Central direction, coordination, and some resource support will be needed.

Users, policymakers, agency officials, librarians, public advocacy groups, and others widely support the vision of a refocused GILS as outlined in this chapter. It is a vision that requires national support. It is a vision that is too important to be ignored. It is a vision that the Administration's efforts to improve the government's ability to provide a range of networked information resources and services clearly support. It is a vision that can be reached.

References

Adams, Margaret O. & Thibodeau, Sharon. G. (1996). The Government information locator service: Origins and potential. *Journal of Government Information*, 23(4), 453-462.

Agencies are generally complying with GILS deadline. (1995, December 1). *Electronic Public Information Newsletter*, 5(12), 180-181.

Aldrich, Duncan M. (1996). Depository libraries, the Internet and the 21st century. *Journal of Government Information*, 23(4), 381-391.

Allerton Institute. (1995). *How we do user-centered design and evaluation of digital libraries: A methodological forum*. 37th Allerton Institute 1995. Graduate School of Library and Information Science University of Illinois at Urbana-Champaign.
[Available:<http://edfu.lis.uiuc.edu/allerton/95/>]

Ambur, Owen. (1996). *Needles in Haystacks: Getting to the Point of Federal Records with Document Metadata and Electronic Document Management Systems*. College Park: University of Maryland, November 23, 1996 DRAFT, [mimeograph].

Baisch, Mary Alice. (1997, March). Reports. Government information locator service conference, November 13-14, 1996. *Documents to the People*, 25(1), 48-50.

Barry, Carol (1994). User-defined relevance criteria: An exploratory study. *Journal of the American Society for Information Science*, 45(3), 149-159.

Bass, Gary D. & Plocher, David. (1991). Finding government information: The federal information locator system (FILS). *Government Information Quarterly*, 8(1), 11-32.

Bertot, John Carlo & McClure, Charles R. (1996a). *Sailor assessment final report: Findings and future sailor development*. Baltimore, MD: Division of Library Development and Services.

Bertot, John Carlo & McClure, Charles R. (1996b). Developing assessment techniques for statewide electronic networks. In S. Hardin (Ed.) *ASIS '96: Proceedings of the 59th ASIS Annual Meeting* (pp. 110-117). Medford, New Jersey: Information Today.

Bertot, John Carlo & McClure, Charles R. (1994). Assessing U.S. government bulletin boards: Problems, policy issues, and recommendations. *Internet Research*, 4(1), 45-63.

Bishop, Ann P. (1995, October). Working towards an understanding of digital library use: A report on the user research efforts of the NSF/ARPA/NASA DLI project. *D-Lib Magazine*. [Available:<http://www.ukoln.ac.uk/dlib/dlib/october95/10bishop.html>].

Bishop, Ann P. & Bishop, Cliff. (1995, Spring). The policy role of user studies. *Serials Review*, 21(1), 17-25.

Buttenfield, Barbara. P. (1995). *User evaluation for the Alexandria Digital Library Project*. [Available:<http://edfu.lis.uiuc.edu/allerton/95/s2/buttenfield.html>].

Chief Information Officers Working Group. (1996, July 31). *Report of the Federal CIO working group*. Washington, DC: Federal Chief Information Officers Working Group.

Christian, Eliot J. (1996a, December). GILS: What is it? Where's it going? *D-Lib Magazine*. [Available:<http://dlib.org/dlib/december96/12christian.html>].

Christian, Eliot J. (1996b). The government information locator service (GILS). *Information Services and Use*, 16(1), 25-42.

Christian, Eliot J. (1995). Government information locator service. *Library Hi-Tech*, 13(1-2), 47-48.

Christian, Eliot J. (1994). Helping the public find information: the U.S. government information locator service (GILS). *Journal of Government Information*, 21(4), 305-314.

Clinton, William J. & Gore, Albert Jr. (1993). *Technology for America's Growth: A New Direction to Build Economic Strength*. Washington, D.C.: Government Printing Office.

Commission on Federal Paperwork. (1977). *The Federal Information Locator System*. Washington, D.C.: Government Printing Office.

Constance, Paul. (1996, June 10). DOD bans rogue web pages. *Government Computer News*, 15(12), 60.

Corbin, Lisa (1996, January). Cyberocracy. *Government Executive*, 28(1), 12-20.

Creswell, John W. (1994). *Research design: Qualitative & quantitative approaches*. Thousand Oaks, CA: Sage Publications.

DiCaterino, Ann & Pardo, Theresa A. (1996, December). *The world wide web as a universal interface to government services*. Albany, NY: Center for Technology in Government, University of Albany. [Available:<http://www.ctg.albany.edu>].

Downing, Thomas A. (1996). Cataloging and locator services for government publications during the transition to an electronic federal depository library program. *Journal of Government Information*, 23(4), 463-470.

Eschenfelder, Kristin, Beachboard, John C., McClure, Charles R., & Wyman, Steve. (1997). Assessing U.S. federal websites. *Government Information Quarterly*, 14(2), 173-190.

Farrell, M. P., Davis, R., Dossett, R., & Baldwin, G. (1996). Electronic initiatives of the federal depository library program. *Journal of Government Information*, 23(4), 393-401.

Federal Information Technology. (1996, July). Executive Order 13011. *Federal Register*, 61(140), 37657-37662.

Finin, Tim. (1997). *UMBC AgentWeb*. [Available:<http://www.cs.umbc.edu/agents/>].

Gellman, Robert. (1996, March 4). GPO shouldn't be the Cinderella of Congress. *Government Computer News*, 15(5), 21..

GILS Conference opens amid perception it is in trouble. (1996, November). *Electronic Public Information Newsletter*, 6(11), 75.

GILS is a finalist for the '96 NII Award. (1996, November). *Electronic Public Information Newsletter*, 6(11), 79.

Glaser, Barney G. & Strauss, Anselm L. (1967). *The discovery of grounded theory: strategies for qualitative research*. New York: Aldine de Gruyter.

Gore, Albert Jr. (1993). *Report of the national performance review: From red tape to results, Creating a government that works better & costs less*. Washington D.C.: Government Printing Office.

Government Printing Office. (1996, March). *Helpful hints for searching federal databases online via GPO Access*. Washington, DC: Government Printing Office.

Hammer, Sebastian & Favaro, John. (1996, March). Z39.50 and the world wide web. *D-Lib Magazine*. [Available:<http://dlib.org/dlib/march96/briefings/03indexdata.html>].

Harrold, Heather. (1997, January 6). 10 web sites take top honors for superlative NII offerings. *Federal Computer Week*, 16(11), 1.

Henderson, Jennifer J. (1997). *Potholes on the information bridge to the 21st century: the second annual OMB Watch Report on the U.S. federal government information locator service (GILS)*. Washington, D. C: OMB Watch.

Henderson, Jennifer N. & McDermott, Patrice. (1995, Summer). Putting government online. *Government Information Insider*, 4, 2-11.

Hernon, Peter, McClure, Charles R. & Relyea, Harold C. (1996). *Federal information policies in the 1990s: Views and perspectives*. Norwood, NJ: Ablex.

Houser, Walter R. (1995, November 27). A pretty GILS is like a memory of all an agency's info. *Government Computer News*, 14(25), 21.

Hsieh-Yee, Ingrid. 1996. Modifying cataloging practice and OCLC infrastructure for effective organization of internet resources. *Proceedings of the OCLC Internet Cataloging Colloquium, January 19, 1996, San Antonio, Texas*. [Available:<http://www.oclc.org/oclc/man/collog/hsieh.htm>].

Information briefs. (1996a, January). *Electronic Public Information Newsletter*, 6(1), 6.

Information briefs. (1996b, March). *Electronic Public Information Newsletter*, 6(3), 6.

Information briefs. (1996c, April). *Electronic Public Information Newsletter*, 6(4), 30.

Information briefs. (1996d, July). *Electronic Public Information Newsletter*, 6(7), 54.

Information briefs. (1996e, October). *Electronic Public Information Newsletter*, 6(9), 70.

Information Infrastructure Task Force. (1994). *The government information locator service (GILS): Report to the information infrastructure task force*, May 2, 1994. Washington, D.C.: Information Infrastructure Task Force. [Available:<http://www.usgs.gov/gils/gils.txt>].

Information Infrastructure Task Force. (1993). *The national information infrastructure: Agenda for action*. Washington, D.C.: Information Infrastructure Task Force. [Available:<gopher://gopher.cni.org:70/00/cnift/p/pub/nni/niagenda>].

Jackson, W. (1996, January 8). Deadline is tight for agencies' directories. (*Government Information Locator Service directories*). *Government Computer News*, 15(1), 6.

Kalil, Thomas. (1995, October). Public policy and the national information infrastructure. *Business Economics*, 30(4), 15-20.

Katzen, Sally. (1996). Welcome from OMB. College Park, MD: NARA II, November 13, 1996 GILS Conference. [Available:<http://www.dtic.dla.mil/gils/carlin.html>].

Koster, Martijn. (1997). *The Web Robots Pages*. [Available:<http://info.webcrawler.com/mak/projects/robots/robots.html>].

Krueger, Richard A. (1988). *Focus groups: A practical guide for applied research*. Newbury Park, CA: Sage Publications.

Lincoln, Yvonne S. & Guba, Egon G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage Publications.

Liu, Jian. (1996). *Understanding WWW search tools*. [Available:<http://www.indiana.edu/~librcsd/search/>].

Love, James. (1994, December 13). Government information locator service. *Tap-Info*. [Available:<http://www.essential.org/listproc/info-policy-notes/0121.html>].

Lynch, Clifford A. (1997a, March). Searching the internet. *Scientific American*, 276(3), 52-56. [Available:<http://www.sciam.com/0397issue/0397lynch.html>].

Lynch, Clifford A. (1997b, April). The Z39.50 information retrieval standard: Part I: A strategic view of its past, present, and future. *D-Lib Magazine*. [Available:<http://dlib.org/dlib/april97/04lynch.html>].

Lynch, Clifford A. (1995, October). Networked information resource discovery: An overview of current issues. *IEEE Journal on Selected Areas in Communications*, 13(8), 1505-1522.

Lynch, Clifford A. (1992). Online searching on the internet: The challenge of information semantics for networked information. In *National Online Meeting*. Medford, NJ: Learned Information.

Lynch, Clifford A., Michelson, Avra, Preston, Cecilia, & Summerhill, Craig A. (1995). *CNI white paper on networked information discovery and retrieval, Chapter 1: The nature of the NIDR challenge* [draft]. [Available: <ftp://ftp.cni.org/CNI/projects/nidr/www/chapter-1.html>].

Machovec, George. (1995, December 1). GILS: the government information locator service. *Online libraries and microcomputers*, 13(12), 1-4.

Mangan, Elizabeth U. (1995). The Making of a standard. *Information Technology and Libraries*, 14(2), 99-110.

Marshall, Catherine & Rossman, Gretchen B. (1995). *Designing qualitative research*. 2nd ed. Thousand Oaks, California: Sage Publications.

Maxwell, Joseph A. (1996). *Qualitative research design: An interactive approach*. Thousand Oaks, CA: Sage Publications.

McClure, Charles R. (1994, Spring). User-based data collection techniques and strategies for evaluating networked information services. *Library Trends*, 42(4), 591-607.

McClure, Charles R. (1991). Planning and evaluation for the networked environment. *EDUCOM Review*, 26 (3-4), 34-37.

McClure, Charles R., Bishop, Ann, Doty, Philip, & Bergeron, Pierrette. (1990). *Federal information inventory/locator systems: From burden to benefit. Final report to the General Services Administration Regulatory Information Service Center and the Office of Management and Budget, Office of Information and Regulatory Affairs*. Syracuse, NY: School of Information Studies. [Available through ERIC Document Reproduction Services ED 326247].

McClure, Charles R., Bishop, Ann, Doty, Philip, & Bergeron, Pierrette. (1991). OMB and the development of a government-wide information inventory/locator system. *Government Information Quarterly*, 8(1), 33-57.

McClure, Charles R. & Lopata, Cynthia. (1996). *Assessing the academic networked environment: Strategies and options*. Washington, D.C.: Coalition for Networked Information.

McClure, Charles R., Moen, William E., & Ryan, Joe. (1992). Design for an internet-based government-wide information locator system. *Electronic Networking: Research, Applications, and Policy*, 2(4), 6-37.

McClure, Charles R., Ryan, Joe & Moen, William E. (1992). *Identifying and describing federal information inventory/locator systems: Design for networked-based locators*. 2 vols. Bethesda, MD: National Audio Visual Center. [Available through ERIC Document Reproduction Services ED 349-031].

McConnell, Bruce W. (1996). New wine in old wineskins: U.S. government information in a networked world. *Journal of Government Information*, 23(3), 217-225.

McKiernan, Gerry. (1996a). *Casting the net: The development of a resource collection for an internet database*. [Available: <http://www.library.ucsb.edu/untangle/mckiernan.html>].

McKiernan, Gerry. (1996b). *Cyberstacks. A Position Paper prepared for the OCLC Internet Cataloging Project Colloquium*. [Available: <http://www.public.iastate.edu/~CYBERSTAC/KS/OCLC.htm>].

Moen, William E. (1995a). The government information locator service: discovering, identifying, and accessing spatial data. In *Geographic information systems and libraries: patrons, maps and spatial information. Proceedings of the 1995 Clinic on Library Applications of Data Processing*, pp. 41-67. Champaign-Urbana, IL: University of Illinois, School of Library and Information Science.

Moen, William E. (1995b). *A guide to the ANSI/NISO Z39.50 protocol: Information retrieval in the information infrastructure*. Bethesda, MD: NISO Press.

Moen, William E. (1994). Information technology standards. A component of federal information policy. *Government Information Quarterly*, 11(4), 357-371.

Moen, William E. and McClure, Charles R. (1997). *Project progress report: An evaluation of the federal government's implementation of the government information locator service (GILS)*. Denton, TX: School of Library and Information Sciences. (Submitted to the U.S. General Services Administration in September 1996 as a deliverable for GSA Contract No. GSA00K96APC2368.)

Moen, William E. and McClure, Charles R. (1996a). *Technical proposal: An evaluation of the federal government's implementation of the government information locator service (GILS)*. Denton, TX: School of Library and Information Sciences. [Available: <http://www-lan.unt.edu/slis/research/gilseval/gilseval.htm>].

Moen, William E. and McClure, Charles R. (1996b). *Work plan: An evaluation of the federal government's implementation of the government information locator service (GILS)*. Denton, TX: School of Library and Information Sciences. (Submitted to the U.S. General Services Administration in September 1996 as a deliverable for GSA Contract No. GSA00K96APC2368.)

Moen, William E. and McClure, Charles R. (1994a). *The government information locator service (GILS): Expanding research and development on the ANSI/NISO Z39.50 information retrieval standard. Final Report*. Gaithersburg, MD: NISO Press.

Moen, William E. and McClure, Charles R. (1994b). The government information locator service: A user-based approach to standards. *StandardView*, 2(2), 86-95.

National Academy of Public Administration. (1996). *Information management performance measures: Developing performance measures and management controls for migration systems, data standards, and process improvement*. Washington, DC: National Academy of Public Administration.

National Archives and Records Administration. (1996). *Ready access to essential evidence: The strategic plan of the national archives and records administration 1997-2007*. Washington, D.C.: National Archives and Records Administration. [Available: <http://www.nara.gov/nara/vision/naraplan.html>].

National Archives and Records Administration. (1995a). *The government information locator service: Guidelines for the preparation of GILS core entries*. Washington, D.C.: National Archives and Records Administration.

National Archives and Records Administration. (1995b). *NARA Bulletin No. 95-03: Government information locator service*. (issued February 16, 1995). Washington, DC: National Archives and Records Administration. [Available: http://gopher.nara.gov:70/0/manage_rs/federal/bulletin/bull95-3.txt].

National Archives and Records Administration. (1990). *Managing electronic records*. Washington, DC: National Archives and Records Administration.

National Historic Publications and Records Commission. (1991). *Research issues in electronic records*. Ellen B. Green (Ed.). St. Paul, MN: Minnesota Historical Society.

National Information Standards Organization. (1995). *ANSI/NISO Z39.50-1995: Information retrieval (Z39.50): Application service definition and protocol specification*. Bethesda, MD: NISO Press. [Available: <http://lcweb.loc.gov/z3950/agency/document.html>].

National Information Standards Organization. (1992). *ANSI/NISO Z39.50-1992, Information Retrieval Application Service Definition and Protocol Specification for Open Systems Interconnection*. Bethesda, MD: NISO Press. [Available: <http://lcweb.loc.gov/z3950/agency/document.html>].

National Institute of Standards and Technology. (1994). Federal information processing standards publication 192, Application profile for the government information locator service (GILS). *Federal Register*, 59 (December 7): 63075-63077.
 [Available:<http://www.dtic.dla.mil/gils/documents/naradoc/fip192.html>].

National Performance Review and the Government Information Technology Services Board (1997, February). *Access America: Reengineering through information technology*.

Network Wizards. (1997). *Internet domain survey, January 1997*.
 [Available:<http://www.nw.com>].

Newby, Gregory B. & Bishop, Ann. P. (1996). Community system users and uses. In S. Hardin (Ed.) *ASIS '96: Proceedings of the 59th ASIS Annual Meeting* (pp. 118-126). Medford, New Jersey: Information Today.

Noonan, Dana. (1996). Making sense of web usage statistics. *The Piper Letter: Uses and Usage*.
 [Available:<http://www.piperinfo.com/piper/pl01/usage.html>].

Okay, John & Williams, Roxanne. (1992). Interagency workshop on public access: A Summary for historical purposes. *Government Information Quarterly*, 10 (2), 237-253.

Office of Management and Budget. (1997a, April 7). *Guidance on developing a handbook for individuals seeking access to public information* [Memorandum]. Washington, DC: Office of Management and Budget.

Office of Management and Budget. (1997b, April 14). *Principles for federal agency use of the world-wide web* [Memorandum (draft)]. Washington, DC: Office of Management and Budget.

Office of Management and Budget. (1996a). *Draft guidelines for agency use of the world-wide web for electronic information collection, access and dissemination, and management*. Washington D.C.: Office of Management and Budget.

Office of Management and Budget. (1996b). *Management of federal information resources (Circular No. A-130)*. Washington, DC: Office of Management and Budget.
 [Available:<http://www.whitehouse.gov/WH/EOP/OMB/html/circulars/a130/a130.html>].

Office of Management and Budget. 1994. *Establishment of a government information locator service. (OMB Bulletin No. 95-01)*. Washington, D.C.: Office of Management and Budget.
 [Available:
<http://www.usgs.gov/gils/omb95-01.html>].

O'Hara, Colleen. (1996a, August 19). GSA prints final guide for X.500 directory. *Federal Computer Week*, 10(24), 14, 19.

O'Hara, Colleen. (1996b, November 18). Report suggests GILS, X. 500 link. *Federal Computer Week*, 10(34), 9.

Olsen, Florence. (1994, April 18). Public locator for government data will make debut on Internet. *Government Computer News*, 13(8), 1.

Olsen, Florence. (1996, April 1). Their goal is to post spatial data for web and GILS users. *Government Computer News*, 15 (7), 1.

Patrick, A. S. (1996). *Services on the information highway: Subjective measures of use and importance from the national capital FreeNet*. Ottawa, Ontario: Government of Canada.
 [Available:
<http://debra.dgbt.doc.ca/services-research/survey/services/>].

Patton, Michael Quinn. (1990). *Qualitative evaluation and research methods*, 2nd ed. Newbury Park, CA: Sage Publications.

Pesachowitz, Alvin. (1992). Interagency conference on public access. *Government Information Quarterly*, 9(2), 187-198.

Phillips, Reed and Carroll, Theresa. (1993). Interagency public access conference: Summary report. *Government Information Quarterly*, 10 (4), 461-476.

Plocher, David. (1996). The paperwork reduction act of 1995: a second chance for information resources management. *Government Information Quarterly*, 13(1), 35-50.

Power, Kevin. (1996, December 9). GILS proves an elusive goal in the federal on-line world. *Government Computer News*, 15(30), 12, 14.

Sikorovsky, Elizabeth. (1996a, January 8). Get the time and expand your mind with web pages. *Federal Computer Week*, 10(1), 16.

Sikorovsky, Elizabeth. (1996b, March 4). GILS breathes new life into web. *Federal Computer Week*, 10(5),

Schneiderman, Ben, Byrd, Don & Croft, W. Bruce. (1997, January). Clarifying search: A user-interface framework for text searches. *D-Lib Magazine* [Available:<http://www.dlib.org/dlib/january97/retrieval/01shneiderman.html>].

Sprehe, J. Timothy. (1997, January 6). EFOIA: mixed bag of access benefits and legal loopholes. *Federal Computer Week*, 11 (1), 15.

Sprehe, J. Timothy. (1996, June 10). Agency home pages getting out of hand. *Federal Computer Week*, 10(14), 10.

Sprehe, J. Timothy. (1996b). U.S. Government Printing Office, no more. *Journal of Government Information*, 23(3), 281-285.

Sprehe, J. Timothy. (1994, April/May). Federal information policy in the Clinton administration's first year. *American Society for Information Science Bulletin*, 20(4), 20-25.

Temin, Thomas R. (1996, March 4). Out of Step: editorial. *Government Computer News*, 15(5), 20.

Thyfault, Mary E. (1994, May 16). Clinton administration steers federal agencies to provide electronic links. *Information Week*, 34.

U.S. Congress. 103rd Congress, 1st Session. 1993. Government Performance and Results Act of 1993: Public Law 103-62, 103rd Congress, 1st Session.

U.S. Congress. 104th Congress, 1st Session. 1995. Paperwork Reduction Act of 1995: Public Law 104-13, 104th Congress, 1st Session.

U.S. Congress. 104th Congress, 2nd Session. 1996. Information Technology Management Reform Act: Public Law 104-106, 104th Congress, 2nd Session.

U.S. Congress. 104th Congress, 2nd Session. 1996. Electronic Freedom of Information Act Amendments of 1996: Public Law 104-231. 104th Congress, 2nd Session.

U.S. National Commission on Library and Information Science. 1992. *Information 2000: Library and information services for the 21st century. Final report of the 1991 White House conference on library and information services*. Washington, D.C.: U.S. National Commission on Library and Information Science.

Van House, N. (1995). UC Berkeley's NSF/ARPA/NASA Digital Libraries Project. *SIGOIS Bulletin*, 16(2), 18-19.

Varon, Elana. (1996, August 26). Worker pushes software for agency accords. *Federal Computer Week*, 10(25), 10.

Weibel, Stuart, Godby, Jean, Miller Eric, & Daniel, Ron. (1995). *OCNL/NCSA Metadata Workshop Report*. [Available:http://www.oclc.org:5046/oclc/research/publications/weibel/metadata/dublin_core_report.html].

Weiss, Peter N. (1997, March 3). Personal communication.

Wilson, Patrick. (1968). *Two kinds of power: An essay in bibliographic control*. Berkeley, CA: University of California Press.

World Wide Web Federal Consortium. (1996). *Guidelines and best Practices*. Washington, DC: World Wide Web Federal Consortium. [Available:<http://www.dtic.mil/staff/cthompson/guidelines/>].

Wyman, Steven K., Beachboard, John C., & McClure, Charles R. (1997). *User and system based quality criteria for evaluating information resources and services available from federal Websites: Final report*. Syracuse, NY: Syracuse University, School of Information Studies.

**APPENDIX A-1
Office of Management and Budget
Bulletin 95-01:
Establishment of Government Information Locator Service**

APPENDIX A-1
Office of Management and Budget Bulletin 95-01:
Establishment of Government Information Locator Service

EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF MANAGEMENT AND BUDGET
Washington, DC 20503

December 7, 1994

OMB BULLETIN NO. 95-01

TO THE HEADS OF EXECUTIVE DEPARTMENTS AND ESTABLISHMENTS

SUBJECT: Establishment of Government Information Locator Service.

1. Purpose. This Bulletin establishes a Government Information Locator Service (GILS), as envisioned in The National Information Infrastructure: Agenda for Action, Information Infrastructure Task Force (September 15, 1993).
2. Authority. OMB Circular No. A-130, "Management of Federal Information Resources," Transmittal Memorandum No. 1, dated June 25, 1993 (58 Fed. Reg. 36068, July 2, 1993), reissued on July 25, 1994 (59 Fed. Reg. 37906).
3. Applicability. This Bulletin applies to all departments and agencies in the Executive Branch. Independent regulatory commissions and agencies are requested to comply.
4. Agency and OMB Responsibilities. Attachment A hereto sets forth the applicable definitions, specifications, implementation schedule, agency responsibilities, and information contacts. GILS will identify public information resources throughout the Federal government, describe the information available in those resources, and provide assistance in obtaining the information. It will also serve as a tool to improve agency electronic records management practices.
5. Termination Date. This Bulletin expires three years from date of issuance.
6. Effective Date. This Bulletin is effective on issuance.

[signed]

Alice M. Rivlin
Director

Attachment A
Establishment of Government Information Locator Service

1. Background. This Bulletin establishes the Government Information Locator Service (GILS) to help the public and agencies locate and access information throughout the U.S. government. It is issued in furtherance of OMB Circular No. A-130, which encourages agencies to ensure public access to government information regardless of form or medium and to establish aids to locating agency information, such as catalogs and directories. The creation of GILS is a goal of The National Information Infrastructure: Agenda for Action which called for the establishment of a "virtual card catalog" of government information holdings. GILS will identify information

resources throughout the Executive Branch, describe the information available, and provide assistance in how to obtain the information. It will improve agencies' abilities to carry out their records management responsibilities and to respond to Freedom of Information Act requests. It will also serve to reduce the information collection burden on the public by making existing information more readily available for sharing among agencies.

GILS will consist of decentralized agency-based information locator records and associated information services. It will use off-the-shelf communications and information technology products and services so that government information can be stored and retrieved in a variety of ways and in a variety of locations.

2. Definitions. As used herein:

"Automated information system" means a discrete set of information resources organized using information technology as defined in OMB Circular No. A-130 for the collection, processing, maintenance, transmission, or dissemination of information which include Federal records as defined in 44 U.S.C. 3301. For purposes of this Bulletin, automated information systems do not include (1) electronic mail and word processing systems, (2) systems the existence of which are specifically authorized under criteria established by an Executive Order to be kept secret in the interest of national defense or foreign policy within the meaning of 5 U.S.C. 552(b)(1), and (3) systems the knowledge of the existence of which would interfere with enforcement proceedings or otherwise be exempt from disclosure under 5 U.S.C. 552(b)(7).

"GILS Core" means a subset of all GILS locator records which describe information resources maintained by Federal agencies, comply with the GILS core elements defined in Federal Information Processing Standards Publication (FIPS Pub.) 192, and are mutually accessible through interconnected electronic network facilities.

"Information dissemination product" means any book, paper, map, machine-readable material, audiovisual production, or other documentary material, regardless of physical form or characteristic, disseminated by an agency to the public. (OMB Circular A-130).

"Locator" means an information resource which identifies other information resources, describes the information available in those resources, and provides assistance in how to obtain the information.

3. Policy.

Section 8(a) of OMB Circular A-130 provides that agencies have a responsibility to "record, preserve and make accessible sufficient information to ensure the management and accountability of agency programs and to protect both legal and financial rights of the Federal Government... provide information to the public consistent with their mission... [and] help the public locate government information maintained by and for the agency." Inventories and finding aids can be an important tool to help other agencies and the public identify information which is available, as well as to help agencies carry out their responsibilities effectively to manage, schedule for disposition and archive their electronic records. Inventories also serve both to increase the efficiency of the dissemination function and to avoid unnecessary burdens of duplicative information collections. The inventories of agency automated information systems and information dissemination products that are reflected in the GILS Core will serve these ends.

Agencies may provide access to their GILS Core locator records either using their own information resources, through an information processing service organization in another agency, through an interagency cooperative effort, or through a contractor. Each agency should establish one or more channels for access to its GILS Core locator records, balancing the goals of facilitating access by the public, assuring appropriate security of government information resources, and minimizing costs to the Government. Direct Internet access to GILS Core locator records should be free of charge, particularly to depository libraries, other libraries, and members of the public with Internet access. Other information dissemination products which include GILS records should be priced in accordance with the provisions of Circular A-130.

Agency GILS are to be established and maintained in accordance with FIPS Pub. 192. As specified in FIPS Pub. 192, the GILS Profile provides the mechanisms for navigating among Federal government locators through specifications given for the GILS Core locator records. Direct users of GILS must be able to use nonproprietary software to access and retrieve information from information sources conforming to FIPS Pub. 192 and the associated GILS Profile. Public domain software that supports access to GILS will be available from the Government Printing Office, the National Technical Information Service, and the Clearinghouse for Networked Information Discovery and Retrieval. GILS will become an integral part of the Federal government's overall information management and dissemination infrastructure, and will ultimately facilitate both identification and direct retrieval of government information. As a first step, agencies should inventory their existing holdings and institute adequate information management practices. To the extent practicable, agency GILS should contain automated links to underlying databases to permit direct access to information identified in the GILS.

4. Responsibilities.

a. All Federal Agencies. The head of each agency should:

- (1) By December 31, 1995, compile an inventory of its 1) automated information systems, 2) Privacy Act systems of records, and 3) locators that together cover all of its information dissemination products. Each such automated information system, Privacy Act system of records, and locator of information dissemination products shall be described by a GILS Core locator record that includes the mandatory GILS Core Elements, and appropriate optional GILS Core Elements as defined in FIPS Pub. 192 and 36 CFR 1228.22(b). Agencies should also supplement the GILS Core Elements with other data elements suitable for specific agency records management and information dissemination needs and objectives. Similar information dissemination products and automated information systems may be identified by a single GILS Core locator record, provided that the locator record clearly identifies the number and scope of items aggregated. Privacy Act systems of records should, however, be identified individually.
- (2) By December 31, 1995, make its initial GILS Core locator records available on-line in a form compliant with FIPS Pub. 192 and the related application profile.
- (3) By June 30, 1996, review the information resources identified in the agency inventory of automated information systems and GILS Core locator records for completeness and to determine the extent to which they include Federal records as defined at 44 U.S.C. 3301. For all Federal records covered by the inventory, the agency shall determine whether they are covered by a records disposition schedule authorized by the Archivist of the United States.
- (4) By December 31, 1996, submit to the Archivist a request for disposition authority proposing schedules for unscheduled records in the information resources described in the GILS Core locator records. The agency should also advise the Archivist if it believes any information resource described in the GILS Core locator records has sufficient historical or other value to warrant continued preservation after the information is no longer needed in the agency.

The inventories of agency automated information systems and information dissemination products that are reflected in the GILS Core should serve as the foundation for developing the records schedules proposed by the agency. When an agency needs to retain different categories of records covered by a GILS Core locator record for different periods of time, the agency should supplement the GILS Core locator record by describing each category. Agencies should cite the applicable disposition authority in the GILS Core element for "supplemental information" for entries that cover records that have been scheduled.

When information dissemination products are part of an on-going series, the agency may submit a proposed records schedule which applies to the entire series. The schedule entry describing such a series may refer to GILS Core locator records to supplement the series description included in the request.

(5) Continually update its inventory and GILS Core locator records as new information dissemination products and automated information systems are identified.

b. Department of Commerce. The Secretary of Commerce should:

- (1) Designate an initial Chair for the Government Information Locator Services Board, established pursuant to section 5, below.
- (2) Maintain FIPS Pub. 192 specifying a GILS Profile with mandatory application for Federal agencies establishing locators and inventories of government information.
- (3) Determine the need for and develop procedures, as appropriate, to identify and validate commercial software packages for compliance with FIPS Pub. 192.

c. National Archives and Records Administration. The Archivist of the United States should:

- (1) Publish guidance and provide, on a reimbursable basis, training to Federal agencies on the development of records inventories, determining records retention needs, and on describing information dissemination products and automated information systems using GILS Core Elements.
- (2) Use GILS entries as an information resource in its records disposition and evaluation programs.
- (3) Cooperate with agencies to reduce reporting burdens and facilitate scheduling of records by accepting GILS data entries when they provide the information required on Standard Form 115, Request for Records Disposition Authority.

d. General Services Administration. The Administrator for General Services should include commercial software packages that implement FIPS Pub. 192 on appropriate Federal Supply Schedules.

e. Interagency Committees. Interagency committees which promote access to and use of Federal information are encouraged to coordinate the efforts of their participating agencies in developing their respective GILS inventories and interagency topical locators when appropriate to their respective missions. Where there is a consensus on the high secondary use value of basic data maintained by multiple agencies, interagency committees should coordinate the development of aggregate information products to meet specific needs identified by communities of interest. These committees include, but are not limited to, the Committees of the National Science and Technology Council, the Federal Geographic Data Committee (FGDC), the Commerce, Energy, NASA, NLM, Defense Information Committee (CENDI), and the Federal Information Resources Management Policy Council (FIRMPOC).

f. Information Processing Service Organizations. Any agency that operates an information processing service organization, as defined in OMB Circular No. A-130, capable of providing on-line access, or other dissemination service, suitable for providing public and interagency access to the GILS, may provide such service for other agencies on a cost reimbursable basis.

5. Government Information Locator Service Board. There is established a Government Information Locator Service Board to evaluate the development and operation of the GILS. Membership on the Board will include representatives of the Director, Office of Management and Budget, the Secretary of Commerce, the Secretary of the Interior, the Archivist of the United States, and the Administrator of General Services. The Public Printer and the Librarian of Congress will be invited to participate as appropriate. The Board may ask the heads of other agencies to designate representatives to serve on the Board or on task forces established by the Board, and should regularly seek comment from State and local governmental entities, interested non-governmental organizations and the public on the operation of the GILS. The Board will prepare and disseminate publicly an annual report that evaluates and recommends enhancements to GILS to meet user information needs, including factors such as accessibility, ease of use, suitability of descriptive language, as well as the accuracy, consistency, timeliness and completeness of coverage.

6. Information contacts.

General policy questions: Peter N. Weiss, Information Policy Branch, Office of Information and Regulatory Affairs, Office of Management and Budget, Room 10236, New Executive Office Building, Washington DC 20503. Telephone: (202) 395-3630.

Records management and archival questions: James J. Hastings, Director, Records Appraisal and Disposition Division, National Archives and Records Administration, Washington, D.C. 20408. Telephone: (301) 713-7096.

Questions regarding FIPS Pub. 192: Shirley Radack, National Institute of Standards and Technology, Gaithersburg MD 20899. Telephone: (301) 975-2833.

General technical questions: Eliot Christian, Information Systems Division, United States Geological Survey, 802 National Center, Reston, VA 22092. Telephone: (703) 648-7245. Electronic mail: echristi@usgs.gov.

7. No Private Right of Action. Nothing in this Bulletin shall be construed to confer a private right of action on any person.

[Available: URL <<http://www.usgs.gov/gils/omb95-01.html>>]

APPENDIX A-2
Paperwork Reduction Act of 1995 (Section on GILS)

APPENDIX A-2

Paperwork Reduction Act of 1995 (Section on GILS)

(U.S. Public Law 44 U.S.C. 3511)

Sec. 3511. Establishment and operation of Government Information Locator Service

- (a) In order to assist agencies and the public in locating information and to promote information sharing and equitable access by the public, the Director shall--
 - (1) cause to be established and maintained a distributed agency-based electronic Government Information Locator Service (hereafter in this section referred to as the 'Service'), which shall identify the major information systems, holdings, and dissemination products of each agency;
 - (2) require each agency to establish and maintain an agency information locator service as a component of, and to support the establishment and operation of the Service;
 - (3) in cooperation with the Archivist of the United States, the Administrator of General Services, the Public Printer, and the Librarian of Congress, establish an interagency committee to advise the Secretary of Commerce on the development of technical standards for the Service to ensure compatibility, promote information sharing, and uniform access by the public;
 - (4) consider public access and other user needs in the establishment and operation of the Service;
 - (5) ensure the security and integrity of the Service, including measures to ensure that only information which is intended to be disclosed to the public is disclosed through the Service; and
 - (6) periodically review the development and effectiveness of the Service and make recommendations for improvement, including other mechanisms for improving public access to Federal agency public information.
- (b) This section shall not apply to operational files as defined by the Central Intelligence Agency Information Act (50 U.S.C. 431 et seq.).

[Available: URL <<http://www.usgs.gov/gils/s244.html>>]

APPENDIX A-3

**National Archives and Records Administration Bulletin 95-03:
Government Information Locator Service**

APPENDIX A-3

National Archives and Records Administration Bulletin 95-03: Government Information Locator Service

National Archives and Records Administration
Washington, DC 20408

NARA BULLETIN NO. 95-3

February 16, 1995

TO: Heads of Federal agencies

SUBJECT: Government Information Locator Service

1. Purpose. This bulletin transmits guidance for U.S. government agencies on describing information dissemination products, automated information systems, and Privacy Act systems using Government Information Locator Service (GILS) Core Elements. It also offers guidance to U.S. government agencies on using GILS as a tool for managing their Federal records.
2. Expiration Date. This bulletin will remain in effect for the duration of OMB Bulletin 95-01.
3. Authority. The Federal Records Act (44 U.S.C. Chapter 33) and implementing regulations (36 CFR Chapter 12) for records management provisions. OMB Bulletin 95-01 for GILS descriptive standards.
4. Background. On December 7, 1994, the Office of Management and Budget issued OMB Bulletin 95-01, Establishment of Government Information Locator Service (GILS). GILS will identify public information resources throughout the Federal government, describe the information available in those resources, and provide assistance in obtaining the information. It will also serve as a tool to improve agency electronic records management practices. GILS will identify and describe Federal information resources and use standard network technology and voluntary, international standards for information search and retrieval to deliver the information to the public.

OMB Bulletin 95-01 directs all Federal agencies to compile an inventory of their automated information systems, Privacy Act systems of records, and locators that together cover all of their information dissemination products and to describe each of these by a GILS Core Locator record. It also directs each agency to determine whether Federal records described in GILS are covered by a records disposition schedule (Standard Form 115) authorized by the Archivist of the United States. Agencies will be able to reduce their reporting requirements by transferring their GILS entries to records schedules or, for systems that are already scheduled, transferring the information on the SF 115 to GILS. Records that are properly described in GILS will be properly described for records schedules, and vice versa.

The OMB Bulletin states that the National Archives and Records Administration (NARA) should publish guidance for Federal agencies on describing information dissemination products and automated information systems using GILS Core Elements. This complements NARA's responsibility to guide agencies on the development of records inventories and in determining records retention needs. The attachment to this NARA Bulletin provides descriptive guidance for agencies to use in completing GILS Core entries.

5. Training. NARA will conduct training courses for agency staff with responsibility for completing the GILS Core entries. NARA will offer these courses on a reimbursable basis. There will be three types of training courses:
 - a. Using the GILS Core Elements. This new one-day course will focus on the descriptive guidance for completing the GILS Core Elements. It will relate this guidance to NARA and OMB records management requirements. The first session of this course will be on March 23, 1995.
 - b. Disposition of Electronic Records. This established two-day course will be modified to add a module on GILS. The module will consider GILS in the context of inventorying and scheduling electronic records. The next session of this course will begin on February 28, 1995.
 - c. Electronic Records Issues. A one-day course on the major issues and problems faced in managing electronic records in Federal agencies, this established course will also add a module for the discussion of GILS. The next meeting of this course will be on March 2, 1995.

For complete information on the schedule and cost of NARA's records management courses, call our office of Records Administration on 301-713-6677. Training information is also available via Internet on gopher.nara.gov.

6. Distribution of Attachment. The attachment to this bulletin is being sent to agencies in two formats. A printed copy and a copy on diskette are being sent to the agency central point of contact who is responsible for distributing NARA bulletins within the agency. In addition, we are providing copies in both formats to agency records officers. The attachment is also available over the Internet on gopher.nara.gov, on the path "Information for Archivists and Records Managers."

TRUDY HUSKAMP PETERSON
Acting Archivist
of the United States

Attachment (see above for location of attachment)

[Available: URL: <<http://gopher.nara.gov:70/0/managers/federal/bulletin/bull95-3.txt>>]

APPENDIX A-4

National Archives and Records Administration Memorandum: Government Information Locator Service and Privacy Act Notices

APPENDIX A-4

National Archives and Records Administration Memorandum: Government Information Locator Service and Privacy Act Notices

** Last update 8/23/95 (lbw) ***

GILS/PRIVACY ACT NOTICES MEMO

This memo from the National Archives and Records Administration's Office of Records Administration has been sent in paper form to Federal agency Records Managers and Information Resource Managers.

MEMORANDUM TO AGENCY RECORDS OFFICERS AND INFORMATION RESOURCE MANAGERS: Government Information Locator Service and Privacy Act Notices

In December 1994 OMB issued Bulletin 95-01, establishing the Government Information Locator Service (GILS). The bulletin lists three categories of information resources that are to be described on GILS: automated information systems, Privacy Act systems, and information dissemination products.

There has been considerable concern expressed by many Federal agencies about the requirement to describe Privacy Act systems on GILS. Agencies already have the responsibility under the Privacy Act to identify their Privacy Act systems through notices in the Federal Register. Agencies consider the responsibility to describe Privacy Act systems on GILS as well as in the Federal Register to be redundant.

We discussed this dual reporting responsibility with OMB and we have agreed that it is an unnecessary burden for agencies to describe by the end of 1995 all of their Privacy Act systems on GILS. As an alternative to describing the systems on GILS, we have entered into an agreement to have the Federal Register Privacy Act notices made available on GPO Access, a Z39.50 compliant server. OMB has agreed that this will fulfill agencies' responsibility for describing their Privacy Act systems on GILS for 1995.

Accordingly, agencies are not required to describe their Privacy Act systems on GILS in 1995. However, we suggest that each agency create one GILS record to indicate to users that their Privacy Act systems are described on GPO Access.

This is an interim solution to the dual reporting requirement. We will continue to pursue an appropriate long-term solution for 1996 and beyond.

Any questions or comments concerning GILS may be directed to our GILS information line. The telephone number is 301-713-7100, ext. 255. Internet inquiries may be sent to GILS@ARCH2.NARA.GOV.

JAMES W. MOORE
Assistant Archivist for
Records Administration

[Available: URL: <<http://gopher.nara.gov:70/0/managers/gils/gilsni.txt>>]

APPENDIX A-5
Report of the December 6, 1995 Meeting of the
Government Information Locator Service (GILS) Board

APPENDIX A-5

Report of the December 6, 1995 Meeting of the Government Information Locator Service (GILS) Board

REPORT OF THE INITIAL MEETING OF THE GOVERNMENT INFORMATION LOCATOR SERVICE BOARD

December 6, 1995

Deputy Secretary of Commerce, David Barram, serving as the Chair, convened the first meeting of the Government Information Locator Service (GILS) Board on December 6, 1995, at 10:00 a.m., in conference room 4830, of the U.S. Department of Commerce's Herbert C. Hoover Building. The Hoover Building is located at 14th & Constitution Ave, NW, Washington, D.C.

He welcomed the attendees and remarked that it had been about one year (December 7, 1994), since he announced GILS on behalf of the Secretary of Commerce, Ronald Brown, the Chair of the Administration's Information Infrastructure Task Force, and also since the signing of the Federal Information Processing Standard 192 which sets out the technical specifications for GILS. He noted that on that same day, Sally Katzen, the Administrator of OMB's Office of Information and Regulatory Affairs (OIRA), announced the signing of OMB Bulletin 95-01 which sets out the policy underpinnings of GILS. GILS received its statutory basis during May 1995 when the President signed into law the Paperwork Reduction Act (PRA) of 1995.

Dep. Sec. Barram thanked Eliot Christian of the Geological Survey and the staffs of the National Archives and Records Administration, the Government Printing Office, and the Interior Department's Geological Survey, for their fine work in assisting the agencies that have been at work developing their GILS implementation. Before introducing the Board, he stated that there would be four informational presentations followed by three business items on the agenda prior to opening the meeting for public discussion.

In his introduction of the Board members, Dep. Sec. Barram again recognized Sally Katzen as the OMB representative. He then welcomed: Governor John Carlin, Archivist of the United States; Mike DiMario, Public Printer; Hiram Davis, Deputy Librarian of Congress; Frank McDonough, representing the General Services Administration; and Robert Lamb, representing the Interior Department. Following the introduction of the Board members, Dep. Sec. Barram introduced Sally Katzen who provided a "charge to the Board".

Sally Katzen thanked David Barram for agreeing to Chair the group. She began by emphasizing the importance of GILS as a valuable component of the National Information Infrastructure which will identify information resources throughout the government and make the information more accessible to the public and agencies. GILS will also improve agencies ability to: carry out their records management responsibilities; respond to Freedom of Information Act requests; and make existing information more readily available for sharing among agencies. She informed the members that the PRA requires the Board to advise the Secretary of Commerce on the development of technical standards for the service. The Board should also consider public access and other user needs, examine the security and integrity of the service, and review its development and effectiveness. She remarked that the first milestone of GILS implementation, scheduled for December 31, 1995, is almost here. This is the date that agencies are required to have their initial GILS inventories available electronically. Most have made progress and some are already operational.

BEST COPY AVAILABLE

PRESENTATIONS

Eliot Christian

Dep. Sec. Barram thanked Sally Katzen and introduced Eliot Christian as the first presenter of four scheduled presentations. Eliot provided an excellent overview of GILS which consisted of a tutorial on locators, a statement of GILS principles, and a report of the current status of GILS in terms of policy and technology. With respect to locators, his presentation covered the topic from its definition, through examples of different types of locators, and finally to how they are created and used. Locators describe information in many forms and can be represented in any media. Eliot stated that the basic principles of GILS are: the adoption of open standards; the support of a diversity of sources; the sensitivity to international languages and standards; the accommodation of copyright, security, and privacy mechanisms; and the extensibility of information extracted from data.

He reported that the Federal government has established strong policies that pertain to GILS or are GILS-related and that the states and foreign countries are actively pursuing GILS and its standards. However, in closing, he cautioned that there is still much to do and that it is still an open question whether or not a world-wide open information infrastructure will be successfully implemented.

Kurt Mulholm

Dep. Sec. Barram thanked Eliot Christian for his presentation and introduced Kurt Mulholm, the Administrator of the Defense Technical Information Center (DTIC), Department of Defense (DOD), to demonstrate DOD's implementation of GILS. Kurt began his presentation by introducing Tammy Borkowski as his associate who would demonstrate the on-line DOD GILS System. Kurt then explained that the DOD GILS system resided on DOD's World Wide Web information service, DefenseLINK, managed by the DOD Office of Public Affairs. He added that Ms. Borkowski, who designed the on-line system, also provided e-mail and floppy disk input capability in recognition that DOD's input providers are disbursed throughout the world. Before turning the presentation over to Tammy, he praised the interagency sharing of knowledge and work effort that has characterized the GILS development environment.

Tammy then conducted the software demonstration on-line via the World Wide Web. Her impressive demonstration consisted of showing how data are inputted to a DefenseLINK Locator Record and how the search and retrieval capabilities of the system operate using a Z39.50 browser add-on for Netscape in Windows.

Keith Belton

Next, Dep. Sec. Barram introduced Mr. Belton from SOLINET, an Internet Service provider. Keith is working with a group of Southeastern states to adapt GILS as a strategic move to enhance their regional information infrastructure. Dep. Sec. Barram remarked that he was especially proud that the GILS concept is being adapted by the States to improve economic development.

Keith described SOLINET's role in the project to develop a Regional Information Service in the southeastern states that will provide access to government information access across jurisdictional boundaries and various levels of governance. He said that at a planning session held in June 1995, participants from 14 states recommended that standards should drive the project and that information for economic development should be the focus of a pilot project. They selected GILS as the most appropriate standard for developing a regional economic database and for its potential for integrating state and Federal locator resources. At the last September 1995 meeting of the Southern Governors Association, the governors decided to make the development of a regional economic database as one of their priority tasks for the coming year.

Keith then presented several issues for consideration by the GILS Board. The Board should consider supporting state efforts to adopt GILS, including encouraging their participation in training programs, and establishing a working group on adoption of GILS at other levels of government. Another issue that he raised was the integration of state and Federal data sources. Dep. Sec. Barram thanked Keith for attending and commented that implementing a regional economic database sounded like an excellent project.

Wayne Kelley

The last presentation was the Government Printing Office (GPO) Pathway Service introduced by Wayne Kelley, Superintendent of Documents, GPO. Mr. Kelley began by saying GILS is a rich and extensible source of government information and various segments of our society will approach GILS from different perspectives. He said the GPO approach is from the point of view of government document librarians and citizens interested in government documents. Moreover, the Pathway Service continues GPO's traditional roles of providing directories and access to government documents and uses electronic technology to achieve significant improvements in service. Wayne then introduced Maggie Parhamovich and Reann Dossett, Internet consultants and key members of the GPO Pathway Service project, to conduct the on-line demonstration.

Reann provided the narrative and Maggie operated the on-line computer interface. Reann explained that the Pathways System is accessed via the on-line GPO Access and will eventually include GILS databases from more than 20 different agencies that have contracted with GPO to put their GILS records on the GPO server. Users can access Pathways by: using the web services; telneting to GPO's WAIS; using a PC-based client; or in the absence of an Internet connection, entering in through the Federal Depository libraries. Reann showed how users can search all agencies' GILS records, or a single agency's database, on specified words or phrases to retrieve the text for review. The search can encompass the full text of the GILS records or be narrowed to concentrate on one or more specific fields. Reann also described the pointer records in the system that provide a direct link to the GILS record inventories of the other Federal departments and independent agencies that are not located on the GPO server. This capability gives users of the Federal Depository Library and others an excellent focal point to access all government GILS records. Reann asked for agency feedback on the Pathways initiative.

BUSINESS ITEMS

Having thanked the GPO presenters for their contribution, Dep. Sec. Barram said there were three business items on the agenda with each presentation taking about five minutes. He then introduced Governor John Carlin, the Archivist of the United States, to discuss the issue of how privacy act systems are described in GILS and then discuss the need for an evaluation of GILS.

Governor John Carlin

Gov. Carlin stated that the issue regarding the requirement for agencies to describe their Privacy Act Systems in GILS and also include them in the Federal Register has been resolved. He said that the National Archives and Records Administration (NARA) has established a partnership with GPO to provide the public with access to Privacy Act Notices on GPO Access, a GILS compliant server. Previously, agencies expressed concern about the responsibility to describe their Privacy Act Systems on GILS when they already had an obligation to describe them in the Federal Register. Gov. Carlin reported that effective December 31, 1995, agencies need only to include a pointer record in their GILS database that points to GPO Access. Also, NARA will distribute a model GILS record to the agencies that contains essential information for the pointer record.

With respect to the second matter, Gov. Carlin emphasized the importance of understanding how well GILS is meeting user information need. Therefore, he proposed that an evaluation be conducted in 1996 that focuses on who has been using GILS, how well their needs have been served, and what, if any modifications are needed to improve service to the public. In addition, he recommended that a final report on the evaluation be submitted to the Board by the end of 1996. In conclusion, he proposed that designated members of the Board be instructed to determine who should conduct the evaluation and how it should be conducted. Gov. Carlin said he would ask his staff to convene the first meeting of the representatives in January 1996.

Dep. Sec. Barram asked for a motion on the proposal for an evaluation of GILS. The motion was made, seconded, and carried by the members.

Michael DiMario

Dep. Sec. Barham introduced Mr. DiMario, the Public Printer, to discuss the work that GPO is doing helping the agencies get started using GILS as well as plans for a GILS point of access.

Mike expressed his pleasure in being able to participate with the Executive Branch in the worthwhile endeavor to improve the dissemination of government information to the public. He remarked that although the demonstration of GPO Access focused on GILS support in the form of high-level agency records, GPO's major thrust is providing GILS compliant servers for mounting GILS for other agencies. He said that 20 agencies had already committed to locate their records on the GILS server in the form of ASCII text. Once resident on GPO Access, agencies can establish a "hot link" from GILS master records to the database(s) on remote server sites. In addition, GPO has provided a World Wide Web service to increase user access to GILS information.

Mike DiMario explained that GPO was able to provide low cost GILS support services to agencies because of the economies of scale resulting from GPO's investment in GPO Access for the Congressional Record and Federal Register. He also expressed his gratitude to Eliot Christian for his advice and guidance to GPO regarding GILS. In addition, he announced that GPO was now making GPO Access free to the public. He said that previously, fee free service was only free to the Federal Depository libraries or a depository gateway.

In conclusion, Mike stated that GPO believes that GILS fills a need for a central, consistent, comprehensive, collection of data about the government's information products and that GPO looked forward to assisting its development and growth.

Dep. Sec. Barram thanked Mike and reiterated the point that agencies should post the address of their GILS' sites on the GILS List to facilitate the establishment of GPO's "point of access" and to make those addresses more easily available to other organizations which may wish to serve as GILS intermediaries.

Steve Hufford

Dep. Sec. Barram then introduced Steve Hufford, EPA, to discuss the work of the existing "GILS Subgroup" and how they can support the wishes of the Board.

Steve began his remarks by saying that the GILS Subgroup is a virtual community of Federal agencies and departments, GILS implementors, commercial interests, information advocacy groups and others and is open to all. He noted that the GILS Subgroup has met monthly for the past year with a focus on sharing information and resolving technical issues related to GILS development.

Steve stated that it was his intent to inform the Board that the Subgroup exists and is a valuable resource for the Board. Accordingly, he suggested a variety of ways that the GILS Subgroup could provide support to the Board. He said that in addition to having the Subgroup undertake Board projects, they might also consider using the Subgroup to: help design and participate in GILS evaluation; bring policy and operational issues to the Boards attention; serve as a channel for communications and outreach; promote coordination with other Federal information locator initiatives; and provide liaison to other governmental level, GILS-related initiatives (states and international).

Dep. Sec. Barram thanked Steve for his presentation and commended the GILS subgroup for the good work that has already been done and expressed his wish that the Subgroup continue its work.

PUBLIC COMMENTS:

Following the scheduled portion of the agenda, Dep. Sec. Barram opened the discussion to members of the public. He asked that each speaker please restrict the length of their remarks so that others might also have sufficient time to speak.

Patrice McDermott

Ms. McDermott, OMB Watch, began her presentation by applauding the progress that agencies have made in identifying and beginning the process of cataloging their information products. She noted, however, that much remains to be done in terms of policy, meaningful access, and public participation. She said that agencies are concentrating on making Web sites available to the public, but few have provided access to the information and data they have collected. Consequently, agencies need to put more information and databases on-line. In addition, agencies need to prepare a comprehensive plan to include identifying what information should be made available to the public and how the information will be accessed. GILS is only one piece of this plan.

Ms. McDermott reminded the audience that many people do not have access to advanced technology and that other means of access, including the Federal depository libraries, need to be employed and supported as effective alternatives. Issues, such as these, need to be addressed by OMB in developing future policy for the access to Federal government information.

Patrice said that when OMB Bulletin 95-01 established the GILS Board, no members of the public, the people using GILS, were included and that this is unacceptable. Consequently, public participation in GILS is minimal. In conclusion, she emphasized that the following questions need to be addressed:

- How will the Federal government provide access to those individuals who do not have access to the World Wide Web?
- How can E-mail documents continue to be excluded from the GILS core records?
- Will the Federal government establish common keyword identifiers for all government information?
- How will the Federal government determine what information citizens want/need?
- How will the Federal government determine user satisfaction with GILS and involve the public and public interest directly in recommending future enhancement and policy directions for GILS?

Ms. McDermott's talk stimulated considerable discussion by the Board members.

Sally Katzen said that although this is a period of scarce resources, the Administration has a strong commitment for supporting GILS and that OMB is doing their utmost to preserve it. She also took issue with Patrice over whether agencies are wrongly investing in establishing Web sites without access to information sources by emphasizing that the ultimate objective is to put agency information on-line. Moreover, agencies cannot put up every piece of information in their inventory immediately. Therefore the Web sites serve both as means of generating public interest and as a logical point of departure for the expansion of the information collections that can be accessed.

Dep. Sec. Barram continued by saying it is his view that this Administration is very oriented for government towards public access and customer satisfaction. He expressed his viewpoint that we must avoid investing resources in making sure that there are multiple ways of getting the information to people who do not have access to it. Rather, we should concentrate on making sure we have access for everybody someplace using the same way of accessing the information.

Patrice McDermott responded to Dep. Sec. Barram's statement by saying that is why OMB Watch was asking for a comprehensive policy.

Dep. Sec. Barram added that he thought most organizations are stronger when they compete with each other for the most satisfied customers.

John Carlin pointed out that the proposal he made for a GILS evaluation was oriented towards the customer, and the Board is not totally out of line on this matter.

Patrice agreed that although the proposed survey of the public and agency uses was a good first step she felt that there should be public involvement in the Board decision-making.

Mike DiMario commented that the Congress, in the GPO Appropriation Act, mandated that he commission a task force to examine electronic delivery and access services to the depository community. This will be an on-going project which requires a report back to Congress this year. Although not stated, there was a strong implication that the task force would include members of the public.

Robert Lamb expressed his uncertainty over the What that will be surveyed in the proposed evaluation. He stated that the evaluation would focus on the Who and How and that Patrice had raised important questions about the data and information distributed on-line. The questions of at what point data becomes information and should be made available to the public does need to be looked at by the Board. However, he said it would be hard to have a comprehensive policy.

Frank McDonough asked Eliot Christian to expand on his voiced concern about the possibility of being unable to achieve the goal of a world-wide, open, information network. He also asked for Eliot's thoughts on where we will end up as a society at the end of the century and what might be the potential barriers.

Eliot Christian responded by saying that the critical thing that he saw occurring was a dangerous trend towards vertically integrated information, content owners and distributors. If a few powerful media or communication giants control the channels of information, the opportunity to have an open network for public access to information might be lost. Rather, the concept behind GILS envisions many ways to find many things. In addition, Eliot noted that commercial organizations are making decisions now which will influence communications and information access five to ten years into the future.

Dep. Sec. Barram presented, by contrast, a more optimistic scenario by saying that there is a powerful feeling in America that there are other ways of doing things. At the turn of the century, things will be dramatically different from the way we think now, communicate, and process information. We will look back to 1995 and wonder what we were worrying about. Addressing Eliot, he commented that although we have to worry a little bit about the future, the history of technology has demonstrated we have always been able to innovate ourselves out of potential or real trouble.

James McDonough

Dep. Sec. Barram recognized the next speaker, Mr. McDonough, the Editor of the Electronic Public Information Newsletter. James stated that his organization has done a survey of the agencies to see how they were doing about implementing GILS. They observed that the majority of agencies planned to put something up on the Internet for GILS. However, they also observed little, if any, coordination between the agency GILS implementors and the agency Public Affairs offices. Public Affairs is the office that interfaces with the public and deals with public policy and what is distributed. In addition, the IRM people, who were implementing GILS, had low level coordinating committees which were not hooked into the agency policy making apparatus. Moreover, they did not get the impression that the agencies viewed GILS as an instrument that they could use to help themselves and the public.

Dep. Sec Barram thanked James for his observations and recognized the next speaker.

Anne A. Hennue

Ms. Hennue, the Associate Director of the Washington Office, American Library Association, introduced herself and responded to Dep. Sec. Barram's earlier reference to making sure we have access for everybody, the nation's have and have not, with regard to information. She emphasized that the Federal Depository Library Program is the

infrastructure that has been in place for over a century to provide that access. It is a partnership between the Federal government and libraries around the country to provide access to government information in a variety of formats. Anne concluded her remarks by emphasizing that the library program is for the public, where they can come for the kind of help that they need in order to use this vast treasure house of information resources.

Dep. Sec. Barram responded by saying he sensed a tremendous rejuvenation in America as the potential of the library back to earlier year as the central point of information. He jokingly added that because of this, he would not mind buying stock in the depository library program if it ever went public.

Anne answered that the program had already gone public, and he was, as a taxpayer, getting a great return on his investment.

CONCLUSION

Dep. Sec. Barram closed the meeting by first asking for any additional comments from the audience and Board members and then stated that the minutes of the meeting will be posted on the GILS List-Server. The minutes of the meeting will serve as the first report of the Board. He thanked the attendees and added that further meetings of the Board would be scheduled as the need arises.

[Available: URL: <<http://www.usgs.gov/gils/board95.txt>>]

**APPENDIX A-6
Statement of Purpose for the
OIW/SIG-LA GILS Subgroup**

APPENDIX A-6

Statement of Purpose for the OIW/SIG-LA GILS Subgroup

Revised Draft

**Statement of Purpose for the
Open Systems Environment Implementors' Workshop
Special Interest Group on Library Applications
Government Information Locator Service Subgroup (OIW/SIG-LA GILS Subgroup)**

Background/History

In 1994, the U.S. Federal government formally began implementing a Government Information Locator Service (GILS) to identify, describe, and provide assistance in obtaining, public information resources. To improve access to information, and in recognition of great changes in information and networking technology, the government had previously convened groups of experts to develop consensus on the appropriate technical approach for GILS. The experts drafted a consensus GILS Profile based on ANSI Z39.50 and other international standards. The profile was approved as an Implementors Agreement and coordinated for approval of the international Open Systems Environment Implementors' Workshop by its Special Interest Group on Library Applications (OIW/SIG-LA). With its standing as an international Applications Profile, the U.S. Federal Government adopted the Implementors Agreement as part of Federal Information Processing Standard (FIPS) 192.

The GILS Subgroup was formed in March of 1995 as part of the OIW/SIG-LA. The GILS Subgroup operated throughout 1995 as an information sharing organization, and contributed valuable support towards successful, initial release of the U.S. Federal GILS in January of 1996. Participation in the GILS Subgroup has always been open to anyone: public and private agencies, organizations, associations, or enterprises, whether profit, non-profit, or academic; government at Federal, state, and local levels, whether U.S., international or other nations; and any other interested parties.

Purpose

The purpose of the GILS Subgroup is to help fully realize the potential of the Government Information Locator Service (GILS) concept, and to promote the development and use of this open systems approach for information search and retrieval. The Subgroup exists to help organizations implement GILS, and also to encourage effective evolution of the GILS standard to meet new uses. To accomplish these purposes, the GILS Subgroup: 1) serves as an open forum for the exchange of ideas on GILS development, use, and refinement, 2) forwards to the OIW/SIG-LA appropriate recommendations for changes to GILS, and 3) promotes sound implementation and broad public awareness of GILS. One emphasis of the Subgroup is to strengthen the U.S. Federal GILS to provide a model and test case for other GILS implementations.

Approach

The Subgroup will create and sustain a community of GILS implementors and users, linking together software developers, information creators, information providers, various levels of government, information advocacy groups, the general public, librarians, records managers, standards bodies, and other interested parties. This community of implementors and users will be enabled through periodic meetings, and through communications on the GILS listserver. All recommendations developed at the periodic meetings will be distributed via the listserver for comment and additional discussion prior to becoming final. Advance notices and brief summaries of all periodic meetings will also be posted to the listserver.

posted on GILS Forum April 4, 1996 -- assumed to be final
[Available: URL: <gopher://gopher.cni.org:70/0R25977-31189-/cniftp/pub/gils/forum/log9604>]

APPENDIX B
List of U.S. Federal GILS Sites as of March 7, 1997

APPENDIX B

List of U.S. Federal GILS Sites as of March 7, 1997

GPO-Brokered GILS	URL(s)
Consumer Product Safety Commission	http://www.access.gpo.gov/su_docs/gils/gils.html
Department of Commerce	http://www.access.gpo.gov/su_docs/gils/gils.html
Department of State	http://www.access.gpo.gov/su_docs/gils/gils.html
Department of Treasury	http://www.access.gpo.gov/su_docs/gils/gils.html
Equal Employment Opportunity Commission	http://www.access.gpo.gov/su_docs/gils/gils.html
Farm Credit Administration	http://www.access.gpo.gov/su_docs/gils/gils.html
Federal Communications Commission	http://www.access.gpo.gov/su_docs/gils/gils.html
Federal Emergency Management Agency	http://www.access.gpo.gov/su_docs/gils/gils.html
Federal Labor Relations Authority	http://www.access.gpo.gov/su_docs/gils/gils.html
Federal Maritime Commission	http://www.access.gpo.gov/su_docs/gils/gils.html
Federal Reserve Board	http://www.access.gpo.gov/su_docs/gils/gils.html
Federal Trade Commission	http://www.access.gpo.gov/su_docs/gils/gils.html
General Services Administration	http://www.access.gpo.gov/su_docs/gils/gils.html
Government Printing Office	http://www.access.gpo.gov/su_docs/gils/gils.html
International Trade Commission	http://www.access.gpo.gov/su_docs/gils/gils.html
Merit Systems Protection Board	http://www.access.gpo.gov/su_docs/gils/gils.html
Office of Government Ethics	http://www.access.gpo.gov/su_docs/gils/gils.html
Office of Management and Budget	http://www.access.gpo.gov/su_docs/gils/gils.html
Office of Personnel Management	http://www.access.gpo.gov/su_docs/gils/gils.html
Overseas Private Investment Corporation	http://www.access.gpo.gov/su_docs/gils/gils.html
Pension Benefit Guaranty Corporation	http://www.access.gpo.gov/su_docs/gils/gils.html
Railroad Retirement Board	http://www.access.gpo.gov/su_docs/gils/gils.html
Securities and Exchange Commission	http://www.access.gpo.gov/su_docs/gils/gils.html
Selective Service System	http://www.access.gpo.gov/su_docs/gils/gils.html
Social Security Administration	http://www.access.gpo.gov/su_docs/gils/gils.html
U.S. Commission on Civil Rights	http://www.access.gpo.gov/su_docs/gils/gils.html
Nuclear Waste Technical Review Board	http://www.access.gpo.gov/su_docs/gils/gils.html
FedWorld-Brokered GILS	
Department of Commerce	http://www.fedworld.gov/gils/
Federal Energy Regulatory Commission	http://www.fedworld.gov/gils/
U.S. Nuclear Regulatory Commission	http://www.fedworld.gov/gils/
Record-Source Hosted GILS	
Department of Agriculture at USDA	http://www.usda.gov/gils/usdagils.htm
Department of Defense at DTIC	http://www.dtic.dla.mil/defenselink/locator/
Department of Energy	http://apollo.osti.gov/osti/gils/gilsintr.html
Department of the Interior at USGS	http://www.usgs.gov/gils/esdd/waisgate.html
Department of Labor	http://bubba.dol.gov/dol/public/dolgils/main.htm
Department of Veterans Affairs	http://www.va.gov/gils/index.htm
Environmental Protection Agency	http://www.epa.gov/earth100/index.html
General Services Administration	http://www.gsa.gov/gils/mosaic.htm
Health and Human Services	http://www.dhhs.gov/progorg/oirm/newhhsgils.htm
Housing and Urban Development	http://www.hud.gov/gils/index.html
National Aeronautics and Space Administration	http://www.sti.nasa.gov/gils/
National Archives and Record Administration	http://www.nara.gov/gils/gils.html
National Labor Relations Board	http://www.doc.gov/nlrb/nlrbloc.html
National Transportation Safety Board	http://www.ntsb.gov/Info/Info.htm
Office of Management and Budget	http://www1.whitehouse.gov/WH/EOP/OMB/html/gils/gils-
Small Business Administration	http://www.sbaonline.sba.gov/gils/
Tennessee Valley Authority	http://www.tva.gov/gils/tvagils_full.htm
United States Postal Service	http://www.usps.gov/GILS/mainlist.html

**APPENDIX C-1
Site Visit Methodology**

APPENDIX C-1

Site Visit Methodology

Table of Contents

1.0. Introduction	1
2.0. Site Visits as a Research Technique in Support of the Project Goals.....	1
3.0. Selection of Site Visits	1
4.0. Selection of Site Visit Participants	3
5.0. Site Visit Demographics.....	3
6.0. Site Visit Data Collection and Data Analysis Activities.....	4
6.1. Content Codes Used in the Database Management Software.....	4
6.1.1. Content-Related Factors.....	4
6.1.2. Critical Success Factors	5
6.1.3. Policy Related Factors	6
6.1.4. Standards	7
6.1.5. Technology-Related Factors	7
6.1.6. Users-Related Factors	7
7.0. Limitations And Conclusion.....	8
7.1. Limitations	8
7.2. Conclusion.....	8
Table C1-1 Site Visit Participants	3

APPENDIX C-1 SITE VISIT METHODOLOGY

1.0. INTRODUCTION

The investigators conducted one-day site visits to four agencies to solicit information about the GILS implementation process at each agency. Investigators held guided interviews with personnel from many administrative and functional areas. Site visits also included a focus group with agency staff, examination of relevant agency documentation, and tours/demonstrations of GILS sites.

2.0. SITE VISITS AS A RESEARCH TECHNIQUE IN SUPPORT OF THE PROJECT GOALS

Site visits provided the investigators with an opportunity for onsite interaction with agency staff involved with GILS implementations. Guided interviews with agency staff focused on individual roles as they related to GILS; on past, present, and future expectations with regard to GILS; on GILS policy issues; and on technology factors which guided agency implementation. Further areas that the investigators explored with agency staff included information about the agency's selection of information resources for GILS record creation, and staff assessment of the data content within GILS records. Investigators probed for staff assessment of critical success factors and barriers encountered during the implementation process. Investigators paid particular attention to staff expectations about users of GILS, both from an internal agency perspective and from a public users perspective. This data collection enabled the investigators to gather information with respect to agency implementation experiences.

3.0. SELECTION OF SITE VISITS

In determining which agencies to visit, the investigators established preliminary criteria for agency site selection. In choosing sites, the investigators wanted a rich mix of the processes, people, programs, interaction, and structures under study (Marshall & Rossman, 1995). The preliminary criteria allowed categorization of agency by implementation effort:

- Agencies implementing GILS servers
- Agencies creating records but not implementing servers
- Agencies which were not initiating any GILS activities (server implementations or record creation).

With these categories in mind, the investigators selected the Environmental Protection Agency (EPA), Government Printing Office (GPO), Department of Defense, Defense Technical Information Center (DTIC), and the Department of Treasury for site visits. Each of these agencies reflected different approaches and responsibilities toward GILS implementations:

- EPA is an example of an agency which had the technical expertise to mount GILS records on its own server and had strong administrative commitment to GILS within the agency. Investigators visited this agency on October 23, 1996.
- The Government Printing Office is an agency which brokers over 2,000 GILS records from 27 agencies. GPO staff created records for GPO resources needing GILS description. The remainder of the records brokered by GPO came from originating agencies that contracted with GPO to make their GILS records available. Investigators visited this agency on November 15, 1996.
- DTIC (now part of Department of Defense's Research and Engineering Directorate under the Under Secretary for Acquisition and Technology) hosts GILS records on its DefenseLINK site, making these records available both on the Internet and on DOD's intranet. Prior to the formal emergence of GILS, DTIC had a need to create and maintain a GILS-like service. Investigators visited this agency on November 15, 1996.
- The Department of Treasury is an example of an agency which mounted its records on GPO's server and chose to create GILS records at the item level. Investigators visited this agency on January 10, 1997.

These four site visits captured different scenarios of agency GILS implementations and enabled the investigators to observe a variety of environments for GILS.

4.0. SELECTION OF SITE VISIT PARTICIPANTS

Site visits provided detailed understanding—from participants' perspectives—of agencies GILS implementation activities and issues. The investigators interviewed as wide a variety of staff at each site as possible to ensure content-rich qualitative data from perspectives that would encompass the full range of an agency's implementation experience. The investigators attempted to systematically identify agency staff for each site visit to provide an agency perspective on all five dimensions of the evaluation framework. Investigators identified the following categories as possible participants in GILS activities in each agency: Administrators/Managers, Policy Makers, Record Managers, IRM staff, GILS Record Creators, Librarians, and Agency Users. In arranging the site visit protocol for each agency, the investigators requested participation by individuals in each of these categories. Investigators also asked key agency contacts to suggest individuals who had been prominently involved in that agency's implementation efforts. The demographics from the site visits reflect the actual group of people who participated in the agency's site visit. The categories of staff interviewed during the site visit depended on that agency's implementation scenario.

5.0. SITE VISIT DEMOGRAPHICS

The investigators talked to a total of 43 site visit participants. Of that number, 40 completed profile sheets. See Appendix D-1 for a copy of the profile sheet. Table C1-1 summarizes the types of staff who participated in the site visits.

Table C1-1
Site Visit Participants

	Treasury	DTIC	EPA	GPO
Total Participants	14	11	10	5
Categories of Staff				
Program Analyst	3			1
Management Analyst	3			1
IRM staff	5	4	4	
Librarian/Archivist	3		2	
Technical Information Specialist		3		
Administrator or Manager		1	1	1
Policy Analyst		1		
Record Manager			2	
Other		2	1	2
Average Yrs/Exp in present position	7	6.5	3.5	3.6
Nature of Work				
CIO/Systems	6	3	2	
Library	3	3	2	
Record Manager	2	0	2	
Project Manager	2	2	1	1
Other IRM	1	1	1	
Public Information Officer		1	1	
Other		1	1	4

6.0. SITE VISIT DATA COLLECTION AND DATA ANALYSIS ACTIVITIES

Participants interviewed during the site visit completed a profile sheet which included quantitative and qualitative questions related to GILS implementations. The profile asked respondents to assess, in a quantitative manner, familiarity with GILS policy documents. Agency staff used a Likert-type scale (from 1-5 in which 1 indicated Very Familiar and 5 indicated Not Familiar) to assess familiarity with GILS policy documents. The profile provided the investigators with assessments of agency staff expectations, training, and lessons learned. The profile sheet also included two questions which asked respondents to identify a favorite online alternative to GILS when trying to locate government information and the reason why this particular alternative was a favorite. These two questions enabled the investigators to identify online sources of government information which were used by agency staff and the personal preferences when agency users wanted to access Federal government information.

Prior to the site visit, the investigators developed interview questions, which were pre-tested by investigators and agency staff. These questions guided the interview process, though the investigators varied from the guided interviews when interviewee responses opened new avenues for data collection. Each site visit had facilitators and recorders who were investigators on the evaluation project. The facilitators and recorders at each session wrote a detailed description of comments made by participants, and an analysis of the issues discussed.

The investigators then created a database from these summaries and used database management software to manage the data collected. The investigators identified coding categories based on the actual data; the evaluation framework for the study sensitized investigators to categories of potential codes. The coding factors represented content found within the narrative summaries. The investigators used coding as a means of analyzing the data obtained from this data collection technique. Once analyzed, the coding scheme provided a data reduction technique for project investigators. As a result of this analysis, investigators were able to query the database for specific incidents of particular factors without losing the ability to focus on the data content from a holistic perspective. The codes used and their definitions are listed below.

6.1. Content Codes Used in the Database Management Software

The following list of codes and their definitions was used to analyze data collected from the site visits.

6.1.1. Content-Related Factors

The following codes identify content issues.

C-DC: Data content for GILS records

This code refers to decisionmaking with regard to data content of GILS records. It includes decisions about which fields to use, record display, granularity levels, and which systems/ product/privacy act systems to describe with GILS records. Issues connected with record creation activity or procedures use another code. References to NARA guidelines use this code if this was the source of the decision making.

C-IR: Effectiveness of Content for Information Retrieval (IR) Purposes

This code refers to the usefulness or value of data in GILS records for either agency staff or users. It includes reference to IR functionality and actual worth of using GILS for information purposes. Doubts or questions about the value of using GILS records to find specific information is included here. Concerns about data display are coded here. Issues about GILS-Lite are coded elsewhere.

C-MAIN: Content Maintenance

This code refers to the maintenance of the data content over time, not the original decisions when the records were created.

C-RC: Record Creation Activity

This code refers to the activity undertaken by agency staff to make GILS records. It includes information about which staff created the records, what procedures staff used, whether the agency made use of contractors to create GILS records, etc. Issues or procedures about the technology support (or lack of support) for record creation are coded elsewhere.

C-RM: Relationship between Record Content and Record Management Purposes

This code is used to bring together record content and record management issues. Policy issues about record management and GILS are coded elsewhere.

6.1.2. Critical Success Factors

The following codes identify critical success factors. Its use here is in the context of key or significant factors relating to agency implementation of GILS.

CSF-\$: Financial Resources

If the agency identified financial resources as an issue or concern when implementing GILS this code is used. This code is used to indicate factors of a financial nature, both the presence or absence of funds.

CSF-AC: Agency Commitment

This code describes the attitudes or behavior exhibited by agency staff which demonstrated a commitment to the GILS initiative. A commitment to public access to government information uses another code.

CSF-AR: Agency Readiness

The key points to this code are pre-existing conditions which enabled the implementation of GILS to happen more effectively or in a more timely manner. Examples include preexisting use of locators, or metadata records. Factors such as technology, money, and administrative management are listed using separate codes.

CSF-B: Benefits

This code is used when there is specific compliance or links between an individual agency's goals and GILS. For a broader commitment to GILS in general, the code *P-B* is used. Doubts as to the usefulness (lack of benefits) when implementing GILS are coded elsewhere but may serve to reflect the antithesis of this code.

CSF-BARR: Barriers

This code refers to barriers or problems identified by agency staff when implementing GILS. Long lists of barriers are coded with *CSF-BARR*; individual records may be coded with a second code if the barrier is in a specific area. By keeping the list of barriers within an agency together, a sense of the whole is maintained along with individual situation. Note: not all problems identified may be of a "critical" nature for an agency but are included here.

CSF-CU: Communication, Feedback with Users

This code refers to the agency's communication with users during the implementation phase. Once the system became functional, ongoing communication is coded *U-COMM*. Examples of the *CSF-CU* might include focus groups.

CSF-MGMT: Effective Management Factors

The key point of this code is the presence or absence of effective managerial planning for GILS. This code would include a pre-existing organizational structure which aided the implementation of GILS or the use of skilled administrators to oversee the initiative. Generally the lack of effective management is coded as a barrier but if the issue includes both positive and negative aspects of management, this code is used.

CSF-SUS: Sustainability

The key point of this code is concerned with issues of sustaining the GILS initiative. Specific issues of record content maintenance used *C-Main*. This code treats broader issues beyond content that an agency must resolve to sustain its GILS initiative. It refers to policies, procedures, and activities involved in sustaining GILS whether implemented or suggested for implementation.

CSF-T: Time factors

This code highlights any concerns from agency staff which regard to time and implementing GILS. Frequently the need to implement GILS in a timely fashion was one of the factors that impacted the overall implementation.

CSF-TE: Technology

The key points connected with this code are the issues or questions about technology that impacted an agency's implementation of GILS. This code is used to describe the effect that an agency's technological infrastructure had on implementing GILS (e.g., whether or not to support its own server). If the technology was a barrier, that code is preferred.

6.1.3. Policy Related Factors

The following codes identify policy issues.

P-95: OMB Bulletin 95-01 policy

This code includes questions or issues connected with the original OMB Bulletin 95-01 and policy issues directly addressed by it.

P-AC: Agency Culture

This code refers to pre-existing agency conditions which impacted the environment in which the GILS implementation occurred. This code provides insights into standard agency ways of doing things and the relationship between the culture within agencies and GILS.

P-B : Benefits of the Policy (awareness, belief in these benefits)

This code is used to describe general benefits that occur to users and agencies as a result of GILS. Specific benefits for an individual agency which motivated that agency to effectively implement GILS are coded as *CSF-B*. Doubts about the existence of benefits are coded elsewhere.

P-CAA: Communication Among Agencies

This code refers to communication which occurs across agencies. It refers to communication between agencies, not within a single agency. This code is preferred for cross-agency issues of a communication nature. Compare with *P-IAG* which is similar but not limited to communication.

P-CB: Costs/Benefits

The key words are costs/benefits and the questioning of the existence of benefits or value to GILS. If someone questions the value of GILS relative to the benefits, it is coded here. Specific barriers to implementation are coded *CSF-BARR* if they occur within a single agency. This code is not limited to a financial assessment of costs and benefits but rather gathers as many instances of references to benefits when coupled with costs.

P-CHAM: GILS Champion

This code is used to address the issue of the need for a GILS champion both within an agency and across agencies.

P-E: Evaluation

This code includes specific reference to the ongoing needs for evaluation or concrete evaluation measures agencies are using or are planning to use when evaluating GILS.

P-FOIA: FOIA

References to GILS and FOIA issues are coded here.

P-F/NS: Future Action or Next Steps in an Agency's GILS program

The key to this code is the identification of policy and/or actions to implement in the immediate future. This code is reserved for lists of multiple suggestions written as a long list of an agency's recommendations. Individual references might use *C-Main*, *CSF-SUS*, or other code. Possible examples: agency activities which are planned in the immediate future; recommendations about policy which are to be implemented by agency administrative staff. Some of the individual items on the list are coded separately by subject. This code preserves an agency-wide view.

P-GD: GILS Definitions

This code is used to describe how people or agency staff define GILS and its purposes. This code captures both articulated definitions and the general confusion about what GILS is supposed to be. General issues about what GILS is, is supposed to be, or confusion about either of the two are coded here.

P-GM: GILS Marketing

This code addresses the concept of GILS marketing and public relations. It is used when someone addressed the issue of marketing either within an agency or across agencies. General recommendations about the need to promote and publicize GILS are coded here.

P-IAG: Interagency GILS Factors

This code speaks to the relationship (or lack of) between agencies or between GPO and other agencies. Any issue which crosses across agencies (except for communication) related to GILS is coded here.

P-IRM: IRM

Specific technical issues within an agency are coded as *CSF-TE*. This code refers to GILS policy issues specifically connected with IRM.

P-M/BC: Metadata and Bibliographic Control

This code speaks to the general concept of metadata and its use as a means that the public uses to find information. It includes the use of bibliographic control. Specific discussion of the needs for standards in the area of record content are coded *S-RC*.

P-PA: Public Access to Government Information

The key point connected with this code is the general issue of public access to government information.

P-RM: Record Management

This code is similar to *C-RM* but broader in that this category speaks to the general relationship between RM and GILS; issues concerned with specific fields or lack of data elements within records are coded as *C-RM*.

6.1.4. Standards

The following codes identify standards issues.

S-A/V: Appreciation and Value of Standards

The key point of this code is an agency staff's awareness of, appreciation for, or value of standards. If the agency staff expressed an attitude that interoperability is a good thing or expressed an interest in promoting standards, this code was used.

S-DC: Standards for Record Content

Many GILS records do not follow uniform guidelines with regard to data content. The need for standards for record content is addressed here.

S-Z: Z39.50

The key word for this code is Z39.50. Specific reference to Z39.50 are coded here.

6.1.5. Technology-Related Factors

The following codes identify technology issues.

T-L: Links between Metadata and Full Text

This code refers to instances in which someone discuss links among GILS records, Web pages, and/or full-text documents. The discussion may include policy regarding these links, or desirability of having links where none exist, or expectations by users for links between text of documents, Web pages, and GILS records.

T-RC: Technology and Record Creation

This code refers to the issues of technology's support of record creation.

T-T: Training

This code refers to instances in which someone discusses the need for training in use of GILS records or describes training activity undertaken by the agency in support of GILS.

T-W: The Web and Its Relationship to GILS

This code refers to the general relationship between GILS records and the Web. If users discuss the relationship between their use of GILS and the Web or if agency staff discuss practices or policies in the agency relating GILS activities and other Web based activities, this code is used. If the discussion emphasis is on links between/among GILS records, or full text documents, or Web pages, the code *T-L* is preferred.

6.1.6. Users-Related Factors

The following codes identify user factors.

U-AS: Agency staff

This code refers to agency staff as users of GILS records. If the agency staff are librarians, the code *U-L* is used. All other types of agency staff as users of GILS use this code.

U-COMM: Communication

The key point of this code is the ongoing channels which an agency establishes to communicate with users. Examples might include email links on GILS records or a user help desk.

U-L: Librarians

This code refers to librarians and their usage of GILS records.

U-P: Public

This code refers to the public as end users of GILS records and their experiences with GILS. If the information primarily refers to effectiveness of IR, it uses another code.

7.0. LIMITATIONS AND CONCLUSION

Site visits offered rich opportunities for data collection in that the investigators were able to evaluate GILS implementations in the context of each agency's own territory. The investigators were able to learn first-hand about GILS implementations at each of the agencies from staff had recent direct experience with the purpose of the evaluation study.

7.1. Limitations

The investigators identified two limitations of the site visits as part of the evaluation process. One limitation involved in the use of site visits was the natural tendency by staff at each of the agencies to put an interpretation on events which would present the agency in as favorable a light as possible. A second limiting factor with this method was the dependence on the availability of particular staff for interviews on the site visit day. If key staff whose insight were considered important in understanding that agency's implementation process were not available on the day of the site visit, the investigators contacted those individuals for subsequent interviews.

7.2. Conclusion

The use of site visits provided the investigators with access to groups of Federal agency staff who were directly involved with GILS. Having access to these individuals, enriched the data collection activity of the evaluation process by providing an extensive understanding of not only what actions were taken by staff but why these actions were taken. The occasion of the site visit also provided agency staff with an opportunity to reflect and analyze their collection actions with regard to GILS implementation.

**APPENDIX C-2
Focus Group Methodology**

APPENDIX C-2

Focus Group Methodology

Table of Contents

1.0. Introduction	1
2.0. Focus Groups as a Research Technique in Support of Project Goals.....	1
3.0. Selection of Focus Groups	1
4.0. Selection of Focus Group Participants	1
5.0. Focus Group Demographics	2
6.0. Focus Group Data Collection and Data Analysis Activities	2
6.1. Content Codes Used for Data Analysis.....	3
6.1.1. GILS Related Codes.....	3
6.1.2. Policy-Related Codes	4
6.1.3. Codes about Users.....	5
7.0. Limitations and Conclusion.....	5
7.1. Limitations.....	5
7.2. Conclusion.....	5
Table C2-1 Focus Group Participant Demographics.....	2
Table C2-2 Focus Group Participant by Type of Work	2

APPENDIX C-2 FOCUS GROUP METHODOLOGY

1.0. INTRODUCTION

The investigators conducted seven focus group sessions designed to elicit perceptions from various stakeholder groups involved with GILS. These focus groups occurred throughout the first six months of the study. A total of 84 participated in these sessions. Table C2-1 summarizes the types of participants at each focus group. These focus groups were in addition to focus groups conducted as part of the agency site visits.

2.0. FOCUS GROUPS AS A RESEARCH TECHNIQUE IN SUPPORT OF PROJECT GOALS

Focus groups brought together GILS stakeholders who shared interests in common themes (e.g., public interest, technology). The focus groups provided opportunities to explore shared beliefs and goals with respect to GILS. The investigators included selected individuals at each focus group to ensure content-rich qualitative data from perspectives that would encompass the range of user and stakeholder beliefs and concerns (Krueger, 1988). The investigators chose participants in the focus groups to meet the information needs represented by the evaluation framework's five dimensions and three perspectives.

3.0. SELECTION OF FOCUS GROUPS

The seven focus groups, dates they were held, and the number of people attending are as follows:

- North Texas Government Document Librarians, 10, (October 31, 1996)
- Public Interest/Public Access, 6, (November 13, 1996)
- State/Local GILS Implementors, 9, (November 13, 1996)
- Future Issues, 11, (November 14, 1996)
- Record Managers, 6, (November 14, 1996)
- Vendor/Technology, 11, (November 13, 1996)
- Small Agency Record Officers and IRM Representatives, 15, (February 13, 1997).

The investigators selected these seven groups to provide opportunities for discussions about GILS implementations with identified categories of stakeholders. Four focus groups represented GILS policy activists and public users, two groups were agency staff, and one group consisted of GILS implementors at governmental agencies other than the Federal level.

4.0. SELECTION OF FOCUS GROUP PARTICIPANTS

The investigators held the first focus group in the North Texas area by interviewing a group of government document librarians at area public and university libraries. The investigators developed focus group questions for this group to assess their awareness and use of GILS. This group of users were intermediaries between the general public and the Federal agencies in that they regularly assisted users in finding government information and would likely be a group for whom GILS would have significant value.

The investigators arranged the next five focus groups at the November 1996 GILS Conference. Conference planners provided the investigators with names of pre-registrants. Investigators analyzed the list of names as to their common characteristics relative to GILS and selected individuals whose responsibilities or expertise matched with information needs of the evaluation study. The investigators invited 12-15 people for each focus group, recognizing that the actual focus group might end up being smaller in number.

The final focus group with the Small Agency Record Officers and IRM Representatives consisted of stakeholders who were involved with GILS implementations at agencies whose staffing resources and information resources were of a significantly smaller size than the site visit agencies. This group's experience with GILS provided the researchers with additional perspectives from the agency and government-wide perspectives of the evaluation framework. The Small Agency Council assisted in arranging this session.

5.0. FOCUS GROUP DEMOGRAPHICS

These focus groups reflected a purposeful sample of GILS stakeholders. The investigators interviewed a total of 84 participants in the focus groups. Of the 84 participants, 68 completed a profile sheet. Tables C2-1 and C2-2 provide summary demographic information of the focus group participants.

Table C2-1
Focus Group Participant Demographics

Name of Focus Group	Number of Participants	Average Years of Experience in Current Position
North Texas Government Documents Librarians	10	5
Public Interest	6	5.5
Future Issues	11	4.9
Record Managers	6	5.5
State/Local	9	4.1
Vendor/Technologists	11	3.4
Small Agency Council	15	6.1

Table C2-2
Focus Group Participant by Type of Work

Type of Work (Job Title)	Number of Participants in All Focus Groups
Administrators/Managers	6
Academics	1
Analysts	4
Archivists	4
Librarians	16
Lobbyist	1
Other	6
Public Affairs	1
Record Managers	14
Systems/Computer Specialists	15

6.0. FOCUS GROUP DATA COLLECTION AND DATA ANALYSIS ACTIVITIES

Participants interviewed during the focus group sessions completed a profile sheet which included quantitative and qualitative questions related to GILS. The profile sheet asked respondents to assess, in a quantitative manner, familiarity with GILS policy documents. Focus group participants used a Likert-type scale (from 1-5 in which 1 indicated Very Familiar and 5 indicated Not Familiar) to assess familiarity with GILS policy documents.

The investigators developed separate profile sheets for each of the focus groups in order to match the information needs of the research project with the various stakeholder groups. The profile provided the investigators with assessments about participant knowledge of GILS policies and attitudes, as well as qualitative information concerning expected user benefits, lessons learned, and perceived barriers or threats to the success of GILS. Included on the profile sheet was a question which asked respondents to identify a favorite online alternative to GILS when trying to locate government information and the reason why this particular alternative was a favorite. These questions enabled the investigators to identify online sources of government information which were used in addition to GILS and would reflect a user-based choice for accessing Federal government information.

Prior to each of the focus groups, the investigators developed interview questions for each session which were pretested by investigators and selected participants. These questions guided the interview process, though the investigators varied from the focus group protocols when interviewee responses opened new avenues for data collection. Focus groups lasted from 45 to 90 minutes. Each focus group had facilitators and recorders. The recorders at each session wrote a detailed description of comments made by participants and an analysis of the issued discussed.

The investigators created a database from these seven summaries and used database management software to manage the data collected. The investigators defined a set of coding categories based on the actual data; the evaluation framework sensitized the investigators to the broader categories. The coding factors represented content found within the narrative summaries. Specific coding categories included categories for GILS Issues, Information Policy Issues, and Users. The investigators used coding as a means of analyzing the data obtained from this data collection technique. Once analyzed, the coding scheme provided a data reduction technique for project researchers. As a result of this analysis, investigators were able to query the database for specific incidents of particular factors without losing the ability to focus on the data content from a holistic perspective.

6.1. Content Codes Used for Data Analysis

There were three general categories of codes developed for the focus group database. These three categories centered around GILS, the policies associated with GILS, and users of GILS.

6.1.1. GILS Related Codes

These codes identify GILS-related issues.

GILS-Alt: Alternatives

This code is used for specific references to Web sites that are alternatives to GILS.

GILS-B: Benefits

This code is used for statements about GILS benefits (as opposed to negatives). It includes stated advantages or positive outcomes as a consequence of GILS.

GILS-Cham: GILS Champion

This code describes someone who is a GILS champion or identifies for the need for such. It includes references to individuals who have acted as product champion. This code may also include reference to the lack of a champion.

GILS-Def: GILS Definition

This code includes comments about how users define GILS and what users think GILS is.

GILS-FDLP: Federal Depository Library Program

This code includes comments about the role/relationship of depository libraries to GILS.

GILS-IR: Information Retrieval

This code includes aspects of GILS which lend themselves to information retrieval. The code includes both functional/mechanical and general design aspects of GILS for IR value. It also includes comments as they pertain to the value of information content within GILS for public use.

GILS-N/P: Negative Aspects and Problems

This code includes comments about GILS which were of problematic nature and included real or perceived negative consequences of GILS implementations.

GILS-P: GILS Purposes

What is the purpose of GILS? This code includes agency and end users understanding in this area. It includes misperceptions of purposes as well as intended GILS purposes. If the comment is primarily directed at the confusion caused by OMB as to the purpose of GILS, the code P-OMB is likely to be preferred. This code is general in scope.

GILS-WEB: GILS Web

The code includes comments about the relationship between the Web and GILS. It includes agency staff and user comments about the value of Web pages and GILS.

6.1.2. Policy-Related Codes

These codes identify policy issues.

P-AG/IAG: Agency Issues and Interagency Issues

This code includes implementation issues across agencies, AILS, agency-related aspects of GILS policy, decentralized/centralized issues about GILS, and the question of who is in charge among the agencies (or lack of interagency leadership). Other codes are used for specific issues such as record management or public access if the comment about agency activity is limited to that topic.

P-F/NS: Future or Next Steps

This code includes comments relating to desired next steps or future actions with respect to GILS and general future needs from a policy perspective. Discussions limited to more narrow improvements are classed under U-Improvements. This code includes both desired future actions and limitations on future action.

P-GILS: GILS Policy

This code includes general comments relating to GILS and a multitude of policy issues too general to be coded elsewhere. It includes both positive and negative aspects of the issues. Generally other codes are preferred if they are more specific.

P-GPO: Government Printing Office

This code includes comments relating to the role of GPO with respect to GILS implementation efforts.

P-INTER: International

This code refers to international uses of GILS.

P-IRM: IRM

This code includes policy aspects of technology, information management, IRM, life cycle management of information, and CIOs.

P-MAR: Marketing

This code includes the need for GILS marketing both to the public and to the agencies. It includes the lack of marketing as a policy problem.

P-NARA: NARA

This code includes specific comments about NARA's role with GILS. General comments about record management are coded elsewhere.

P-OMB: GILS and OMB

This code include specific references to OMB's actions or lack of actions with respect to GILS. Anger directed at OMB and confusion about GILS purposes directed at OMB are coded here.

P-PA: Public Access

This code includes general information about the role of the public's access to information and the use of the phrase public access as an issue. It includes both negative and positive support for public access to agency information.

P-RC/RC: Record Content And Record Creation

This code includes comments both about record content and record creation, comments about the nature of information data within GILS records, discussion about levels of granularity, decisions regarding which records to create in GILS, and discussions about who handles record creation. It includes issues connected with the use of metadata records unless the emphasis is on metadata as a standard.

P-RM: Records Management

This codes comments about policy with respect to GILS and record management.

P-STAND: Standards

This codes includes issues about both technology and metadata standards.

P-ST/LO: State and Local

This codes includes comments predominantly made in the state and local GILS implementors focus group. It provides an review of GILS implementations at these levels and attitudes about implementing GILS among this group of stakeholders.

6.1.3. Codes about Users

These codes identify user issues among the focus groups.

U-EVAL: Users—Evaluation

This codes includes judgmental views of GILS made by users or judgmental views of GILS as GILS pertains to users. It does not include descriptive views.

U-EXP: Users—Experience

This code includes descriptive information about the experiential use of GILS. It is not evaluative, and may make mention of time or number of GILS records examined

U-IMPROVE: Users—Suggested Improvements

This code includes specific suggestions as to needed improvements to GILS. These references are more specific than the P-F/NS code and pertain to users.

U-TECH: Users—Technology Aspects

This code concerns user technology needs when accessing GILS.

7.0. LIMITATIONS AND CONCLUSION

Because of the role that stakeholders hold within the context of GILS, participants at the focus groups actively aided the investigators by providing insight and analysis appropriate for data collection. An advantage of using focus groups for this type of evaluation is that individuals are generally interested and pleased to provide their opinions.

7.1. Limitations

The investigators identified two limitations connected with the use of focus groups in this project. Some participants in focus group sessions asserted their positions emphatically causing the focus group facilitators difficulty in balancing the contributions made by such individuals compared with others in the group. The goal of the sessions was a group discussion rather than a discussion focused on the interests of one or two individuals. Secondly, the focus group facilitators were aware that existing political or administrative relationships which existed between individuals outside the focus groups caused some participants to self-censor remarks given the presence of others in the group

7.2. Conclusions

The use of focus groups provided the investigators with access to groups of individuals who were key stakeholders in GILS. The investigators benefited from the opportunity to reach such a significant number of these people at the GILS Conference. The focus groups complemented the site visits conducted during the study in that the site visits predominantly represented opportunities to gain information about agency implementation while the focus group sessions provided opportunities to learn about GILS from users, policy makers, vendors, state and local GILS implementors, and librarians.

APPENDIX C-3
Gils Conference Survey Methodology

APPENDIX C-3

Gils Conference Survey Methodology

Table of Contents

1.0. Introduction	1
2.0 A Survey as a Research Technique in Support of the Project Goals	1
3.0. Conference Survey Demographics	1
4.0. Conference Survey Questions	2
5.0. Conference Survey Data Collection and Data Analysis Activities	2
6.0. Limitations and Conclusion.....	3
6.1. Limitations.....	3
6.2. Conclusion	3
Table C3-1 Conference Survey Demographics: Respondents' Involvement with GILS	1
Table C3-2 Conference Survey Demographics: Fields In Which Respondents Worked.....	2

APPENDIX C-3

GILS CONFERENCE SURVEY METHODOLOGY

1.0. INTRODUCTION

On November 13 and 14, 1996, the National Archives and Records Administration (NARA) hosted the 1996 GILS Conference. Over 200 people attended the conference. The investigators developed a survey instrument to be completed by participants at the GILS Conference. The survey offered the opportunity to collect information to gauge respondents knowledge and awareness of GILS related policy and technologies. The investigators that the self-selected participants at the Conference represented primarily Federal agency staff who were knowledgeable and interested in GILS. Respondents completed the survey early in the first day of the conference so that investigators could assess existing knowledge about GILS rather than knowledge gained from the Conference. A total of 181 GILS Conference participants completed the survey. Appendix D-3 contains as copy of the survey instrument.

2.0 A SURVEY AS A RESEARCH TECHNIQUE IN SUPPORT OF THE PROJECT GOALS

This research technique solicited information from a set of people who were knowledgeable and or interested in GILS. Although not sampling the entire population of those knowledgeable and or interested in GILS, the survey reflected a purposeful sample, designed to meet the information needs of the evaluation framework's five dimensions and three perspectives. The investigators designed the survey to capture user perceptions and expectations concerning themes and issues pertaining to GILS.

3.0. CONFERENCE SURVEY DEMOGRAPHICS

Of the 181 respondents, 78% worked in a Federal agency, 9% worked for a state or local government, and the 13% identified work settings that included private non-profit and for-profit organizations. The respondents involvement with GILS included a number of different responsibility areas and reflected different areas of expertise. Tables C3-1 and C3-2 summarize respondent demographic and other information.

Table C3-1
Conference Survey Demographics:
Respondents' Involvement with GILS

Area of Involvement With GILS	N	%
Implementors	81	39
GILS Record Creator	46	22
User	39	19
Policymaker	19	9
Technical Standards Developer	5	32
Information Reseller	4	2
Integration/Interoperability	2	1
Trainer/User Support	2	1
Other	8	4
Total	*206	**99

Respondents marked multiple choices so N exceeds the 181 completed surveys.

**Total does not equal 100% due to rounding

Table C3-2
Conference Survey Demographics:
Fields In Which Respondents Worked

Setting	N	%
Records Management	67	34
Information Resources Management	35	18
Library/Information Center	29	15
Public Information	19	10
Program Office/Project Management	10	5
Computer Systems	8	4
Archives	7	4
Chief Information Office	6	3
Legal/Legislative	4	2
Software Developer	3	2
Sales/Marketing	3	2
Research and Development	1	1
Procurement/Contracting	0	0
Other	8	4
TOTAL	*200	**104

*Respondents marked multiple choice so N exceeds the 181 completed surveys.

**Total does not equal 100% due to rounding.

4.0. CONFERENCE SURVEY QUESTIONS

The survey asked respondents to assess, in a quantitative manner, key GILS policy issues. Survey respondents used a Likert-type scale, (from 1-5 in which 1 indicated Very Familiar and 5 indicated Not Familiar) to assess familiarity with GILS policy documents. A second Likert-type scale (from 1-5 in which 1 indicated Strongly Agree and 5 indicated Strongly Disagree and a 6th category of "Don't Know") asked respondents to express opinions about GILS. The investigators included 17 questions concerned with key issues related to GILS.

Included in the conference survey were two questions which asked respondents to identify a favorite online alternative to GILS when locating government information. These questions allowed the investigators to learn of user-based choices for accessing government information and the reason why this choice was a favorite. The survey also included a qualitative question in which respondents could state any comments, issues or topics of interest to the investigators.

5.0. CONFERENCE SURVEY DATA COLLECTION AND DATA ANALYSIS ACTIVITIES

The investigators pre-tested the survey instrument with members of the evaluation team and with project advisory group members prior to its use at the GILS conference. The investigators distributed the Conference survey early during the first day of the GILS Conference. The investigators designed the opportunity to survey respondents to occur prior to any conference panels or presentations in order to prevent these events from modifying participant opinions about GILS. Attendees completed the one-page survey and returned the forms to the investigators.

Investigators used a database program to enter and store data collected from the survey for subsequent manipulation. To verify accuracy in data entry, the investigators randomly selected 20 surveys (approximately 10% of the total) and reviewed their corresponding database entries. Descriptive statistics were the primary output from the survey questions. Output included mean and mode for each question.

6.0. LIMITATIONS AND CONCLUSIONS

Using a survey early in the study provided the investigators with a mechanism to gauge key issues for additional examination through subsequent data collection activities.

6.1. Limitations

The investigators were aware of the limitations in using a survey as a data collection technique for evaluation purposes. As with any survey, each respondent may interpret questions on the survey differently from others respondents, generating data that may be skewed due to user perceptions. The study team discussed at length, during its pretest of the survey, how to minimize this outcome. The investigators felt that the incidence of this problem had been minimized by rewording any question that appeared to have ambiguity in its meaning but recognize that this factor can never be completely eliminated.

6.2. Conclusion

The thoroughness with which respondents answered the survey questions enabled the investigators to use this data collection technique effectively. The high response rate (181 out of 200+ registrants) provided the investigators with a broad-based survey of user knowledge and opinions about GILS related policy and implementation experiences.

APPENDIX C-4
Record Content Analysis Methodology

APPENDIX C-4

Record Content Analysis Methodology

Table of Contents

1.0. Introduction	1
2.0. Method Overview.....	1
3.0. Objectives	1
4.0. Context Within The Evaluation Framework.....	2
5.0. Data Collection and Analysis	2
5.1. Survey of GILS Universe.....	2
5.2. Development of Analysis Criteria.....	4
5.2.1. Issues in Developing Record Content Aggregation Criteria	5
5.3. Content Analysis of Sampled Records.....	6
6.0. Method Limitations and Recommendations to Future Researchers.....	8
7.0. Conclusion.....	8
Table C4-1 Record Content Analysis Sample Population.....	3
Table C4-2 Record Content Analysis Criteria.....	4
Table C4-3 Aggregation Semantics.....	6
Table C4-4 Information Object Semantics.....	7

APPENDIX C-4 RECORD CONTENT ANALYSIS METHODOLOGY

1.0. INTRODUCTION

Moen and McClure, in *The Government Information Locator Service (GILS): Expanding Research and Development on the ANSI/ISO Z39.50 Information Retrieval Standard: Final Report* (1994, p. 30) noted "an important factor in the overall utility of the GILS will be the quality of the data in GILS records. Quality criteria will include accuracy, consistency, completeness, and currency. In order to encourage the creation of high quality information that will populate GILS servers, the development of written guidelines for creating GILS records is essential." This direction, *The government information locator service: Guidelines for the Preparation of GILS Core Entries* (National Archives and Record Administration, 1995a) is available electronically from the National Archives gopher at <gopher.nara.gov> under "Information for Archivists and Records Managers/GILS Guidance," or from <URL: <http://www.nara.gov:70/1/managers/gils>>. In addition, *Federal information processing standards publication 192, Application Profile for the Government Information Locator Service (GILS)* (National Institute for Standards and Technology, 1994) provides other quality-related direction such as preferred order of display for record elements as well as their definitions.

Content analysis of GILS records served three purposes: to assess records' quality in terms of completeness and accuracy; to explore the relationship of selected characteristics of records and serviceability in networked information discovery and retrieval (NIDR); and to develop recommendations for future application or adaptation of the method.

More than 3500 instances of metadata were evaluated for incidence and/or content, and entered into a database for coding and analysis. In addition, the evaluators maintained a log of lessons learned and areas for further research (see Appendix E-2 Record Content Analysis Findings, Discussion, and Recommendations) that may be utilized by system developers, specification and procedures writers, and people with direct responsibility for GILS record quality.

2.0. METHOD OVERVIEW

The analysis comprised in two phases: Phase 1 involved examination of a pool of 83 records from 42 agencies' GILS retrieved deliberately to represent a range of information resource types (e.g., databases, catalogs, records systems). These records served as the basis for developing and operationalizing a set of more than 50 qualitative and quantitative evaluative criteria that included records' format, aggregation, media representation, and descriptiveness. Descriptiveness was defined as the incidence of utilization and content (value) attributes for all mandatory and selected optional elements and subelements as specified by *FIPS Pub. 192 Annex E-GILS Core Elements* and the *NARA Guidelines*. In Phase 2, these criteria were systematically applied to a set of 83 records randomly retrieved January 13 and 14, 1997, from 42 agencies' GILS.

The following paragraphs present information concerning the record content analysis objectives, the context of the analysis within the overall evaluation framework, data collection and analysis, method limitations, lessons learned, and recommendations.

3.0. OBJECTIVES

This analysis attempted to describe the "quality" of GILS records in terms of character or attributes rather than strict conformance to specifications. The latter, which constitutes an audit, would require a greater level of operational detail than current policy and standards provide and is a technique better suited to a more mature information service. The following objectives guided the current examination of GILS records. Where adherence to published direction was relevant, *FIPS Pub. 192 Annex E* definitions, as reproduced and supplemented by usage guidelines and examples in the *NARA Guidelines*, served as the basis for evaluation:

1. To assess the accuracy of GILS records in terms of errors in format and spelling
2. To gauge and compare the relative "completeness" or level of description of GILS records
 - Number of elements per record ("blank" vs. populated)

- Utilization and values of both mandatory and selected optional elements
- 3. To characterize a general profile of GILS product in terms of record types, aggregation levels, and containers (dissemination media)
- 4. To evaluate records' serviceability
 - Factors affecting NIDR
 - User convenience
 - Aesthetics and readability
 - Relevance judgment.

The quantitative and qualitative assessments, respectively, of the constitution and properties of sampled records provided data meeting these objectives.

4.0. CONTEXT WITHIN THE EVALUATION FRAMEWORK

As with the other methods comprising this user-oriented evaluation of GILS implementations, the record content analysis both was informed by and served to inform other data collection and instrument development activities in the study. Presentations and panel discussions at the 1996 GILS Conference and focus groups with various user communities highlighted recurring issues surrounding the content of GILS records, such as the level of resource aggregation, suitability of metadata elements, consistency, and quality of presentation. In turn, as discussed in Appendix E-2 Record Content Analysis Findings, Discussion, and Recommendations, the record content analysis proved invaluable in developing a user-assessment script that would both isolate GILS "quality" from that of the user interface or search engine and present realistic information retrieval encounters.

5.0. DATA COLLECTION AND ANALYSIS

Data collection and analysis were performed as described in the following paragraphs using the tool presented in Appendix D-4 Record Content Analysis Instrument as constructed in a Microsoft Access® database and Microsoft Excel® spreadsheets. Two surveying activities were prerequisite to the analysis of record content: a determination of the GILS universe to optimize the breadth of the sample and a review of planned (i.e., per the NARA *Guidelines*) vs. actual record characteristics to inform development of analysis criteria.

5.1. Survey of GILS Universe

To provide the broadest possible base for record selection, the investigators first determined the universe of GILS implementations. This was accomplished through various means:

- Verbal or written mention during the 1996 GILS Conference presentations and in handouts and survey responses, respectively
- Linking from the White House website's "President's Cabinet" <http://www.whitehouse.gov/WH/Cabinet/html/cabinet_links-plain.html> and "Federal Agencies and Commissions" <http://www.whitehouse.gov/WH/Independent_Agencies/html/independent_links-plain.html> to agency homepages, which, in turn, linked in some cases to FedWorld GILS <<http://fedworld.gov/gils>>
- WWW searches by means of Alta Vista <<http://www.altavista.digital.com>> and Lycos <<http://www.lycos.com>> search engines for Executive department and agency names
 - As delineated in the 1996-97 *Government Manual* via the Government Printing Office (GPO) GPO Access http://www.access.gpo.gov/su_docs/aces/aaces002.html
 - As comprising the Chief Information Officer Council as specified in Executive Order 13011 of July 16, 1996 "Federal Information Technology" (<http://www.gsa.gov/irms/ka/regs/exo13011/exo13011.htm>)
- WWW searches by means of Alta Vista and Lycos search engines for "GILS" and for "government information locator service"
- GPO Access GILS server.

Results of this effort, completed on December 31, 1996, are shown in below in Table C4-1 Record Content Analysis Sample Population with two additional agencies identified for sampling in Phase 2 of the record content analysis.

Table C4-1
Record Content Analysis Sample Population

Consumer Product Safety Commission
Department Of Agriculture
Department Of Commerce
Department Of Defense
Department Of Energy
Department Of Health And Human Services
Department Of Housing And Urban Development
Department Of Interior
Department Of Labor
Department Of State
Department Of Treasury
Environmental Protection Agency
Equal Employment Opportunity Commission
Farm Credit Administration
Federal Communications Commission
Federal Emergency Management Agency
Federal Energy Regulatory Commission
Federal Labor Relations Authority
Federal Maritime Commission
Federal Reserve Board
Federal Trade Commission
General Services Administration
Government Printing Office
International Trade Commission
Merit Systems Protection Board
National Aeronautics And Space Administration
National Archives And Records Administration
National Transportation Safety Board
Nuclear Regulatory Commission
Nuclear Waste Technical Review Board
Office Of Government Ethics
Office Of Management And Budget
Office Of Personnel Management
Overseas Private Investment Corporation
Pension Benefit Garanty Corporation
Railroad Retirement Board
Securities And Exchange Commission
Selective Service System
Small Business Administration
Social Security Administration
U.S. Commission On Civil Rights
U.S. Postal Service
Total=42

5.2. Development of Analysis Criteria

The second activity to prepare for a systematic analysis of GILS record content was the creation of criteria to satisfy the study objectives. This was accomplished by examining a set of two records retrieved from each identified agency GILS. These records—retrieved by use of search terms including “system,” “database,” “manual,” the agency acronym, subject-oriented single words—were selected to represent a variety of file sizes, formats, and content types.

These records were studied and compared to produce the assessment categories shown in Table C4-2 Record Content Analysis Criteria. (Appendix D-4 Record Content Analysis Instrument presents a table of the database fields, possible values, and coding notes that was constructed to record data.)

**Table C4-2
Record Content Analysis Criteria**

Accuracy

- Format and Formatting Errors
- Spelling And Typographical Errors

Completeness

- Number Of Elements Per Record
- Practice Of Presenting “Blank” (Nonpopulated But Displayed) Elements
- Utilization And Selected Characteristics Of “Mandatory” Elements
 - Title
 - Originator
 - Local Subject Index
 - Abstract
 - Purpose
 - Agency Program
 - Availability-Distributor
 - Availability-Order Process
 - Sources Of Data
 - Access Constraints
 - Use Constraints
 - Point Of Contact
 - Schedule Number
 - Control Identifier
 - Record Source
 - Date Of Last Modification
- Utilization And Characteristics Of Selected “Optional” Elements
 - Controlled Vocabulary-Index
 - Terms-Controlled
 - Controlled Vocabulary-Thesaurus
 - Local Subject Index
 - Availability-Resource Description
 - Methodology

Profile

- Record Types (AIS, locator, Privacy Act system)
- Record Aggregation (See Table C4-3 Aggregation Semantics and discussion)
- Objects Represented (see Table C4-4 Information Object Semantics)
- Containers (Dissemination Media)
 - Broadcast (Radio/TV)
 - CD-ROM
 - Dialup
 - Email
 - Fax
 - Ftp Site
 - Gopher Site
 - Listserv
 - Microform
 - Multiple
 - Print
 - Video
 - Voice
 - Web

Serviceability

- Capitalization
- Citation Of Legislation
- Definition Of Acronyms
- Element Display Order
- Fielded-Search Option
- File Formats
- Hypertext
- Indentation
- Locally Defined Elements

5.2.1. Issues in Developing Record Content Aggregation Criteria

The following definitions served as an initial starting point for operationalizing the phenomenon of aggregation:

AGGREGATION: the degree to which two or more separate parts have been brought together without changing their function or producing any result other than the sum of the operation of the parts.

GRANULATION: the degree to which two or more separate parts of a whole are distinguishable within that whole.

It became apparent during review of the Phase 1 sample that the above definitions are unsuitable for application to GILS records. For example, a record describing a publicly-accessible enterprise-wide AIS whose function is to track information output of four discrete, functionally dedicated, not publicly accessible micro-AISs could be labeled a "highly aggregated" record in that it "rolls up" other potential records. But, should the record include a description of each "grain" (microsystem) it embraces, one would be tempted to code it "low granularity" (subparts are distinguishable).

Another, more concrete, example of the problem of characterizing aggregation of information resources would be *The Federal Register* in digital (databased) or paper print format. This one record describes one "discrete" publication, but that publication aggregates myriad standalone information objects that, in print, are highly granular to the initiated user but in database form (digital format) are less distinguishable.

Another, more concrete, example of the problem of characterizing aggregation of information resources would be *The Federal Register* in digital (databased) or paper print format. This one record describes one "discrete" publication, but that publication aggregates myriad standalone information objects that, in print, are highly granular to the initiated user but in database form (digital format) are less distinguishable.

In short, the attribute of "aggregation" is discernible only to the degree that the GILS record presents an explicit enumeration of "granules" or aggregated parts—whether those parts are:

- book chapters,
- database fields,
- Web page titles, or
- Privacy Act records,

which some will argue is too granular, or they are:

- individual reporting systems of enterprise-wide AIS,
- titles within a videotape series, or
- memoranda within a "file,"

which some will argue should be distinguishable.

Application of definitions of aggregation and granularity imply a knowledge of component-level and collective functionalities that the investigators, and, by proxy, a GILS user, lack and which may be gained only through examination of the object. In a physical library, users of a card catalog, subject bibliography, or other metadata-based tools are accustomed to retrieving and scanning resources' object-peculiar "primary" metadata (e.g., tables of content, graphics, and back-of-the-book indexes) as required to determine whether "granules" might satisfy their information need; in GILS, where often information resources cannot be examined and thus their "operation" is unknown, the concept of simply "pointing" to an aggregated "locator" may be insufficient in that the aggregation "produces no result other than the sum of the operation of the parts."

Nonetheless, because record and resource aggregation was identified as a recurring theme during other data collection activities of the study, investigator's adopted the operational definitions of aggregation coding scheme shown in Table C4-3 Aggregation Semantics to characterize the phenomenon. To supplement the limited value returned from assigning aggregation-level coding, investigators incorporated the criterion of "information object" as defined in Table C4-4 as well. Appendix E-2 Record Content Analysis Findings, Discussion, and Recommendations offers additional interpretation of the utility of these measures relative to aggregation and resource description.

Table C4-3
Aggregation Semantics

Code	Operational Definition	Examples
Record Aggregates Objects	GILS record, by virtue of its creation, collects discrete information resources that record content indicates would not have otherwise been collected or aggregated. Assigned in the absence of clues within the record that the represented objects were heretofore packaged <i>as this collection</i> to optimize information discovery and retrieval.	<ul style="list-style-type: none"> • Privacy Act Systems compilation files • press releases forms
Aggregated Object Represented	GILS record represents an <i>a priori</i> or purposeful collection of information resources—e.g., woodpecker database or agency website. GILS record represents an object that collects, or comprises, two or more discrete information objects, and that represents a collection of standalone information files or products packaged together on the basis of a common theme or subject for functional convenience.	<ul style="list-style-type: none"> • CD-ROM of regulations • System that compiles Privacy Act records • job line of open requisitions
Discrete Object Represented	GILS record describes a standalone document-level entity that does not meet the criteria for “object aggregates metadata” below.	<ul style="list-style-type: none"> • annual report • videotape
Object Aggregates Metadata	GILS record describes a pre-existing metadata collection, or “locator,” as an information resource.	<ul style="list-style-type: none"> • directory • catalog • index • log

5.3. Content Analysis of Sampled Records

As of early January 1997, 42 agencies' GILS had been discovered by procedures identified in Section 5.2 Survey of the GILS Universe. The 83 sampled records, selected as described in the next paragraph, resided in three broad “host” categories: GPO (61% of the sample), record sources (34%), and FedWorld (5%). 93% of sampled records resided on a WAIS or Z39.50-compliant server, with the remaining on an HTTP server containing standalone HTML files of GILS records. (Note: since the time period of analysis, FedWorld and GPO have mounted record-source hosted GILS and those hosted by one another, and at least one HTTP-based GILS has migrated to WAIS).

The record content analysis *per se* first involved selection of GILS records from the known GILS universe (see Table C4-1 Record Content Analysis Sample Population) in one of two ways. For GILS featuring a search engine (i.e., residing on an information retrieval-based platform such as WAIS or Z39.50-compliant server or including a site-resident search engine), the investigator retrieved the first and last “hits” resulting from a “full-text” query of the agency acronym (using the default “number of records to return”). For GILS on which this was not possible (i.e., those mounted on a web server of HTML files that present only a picklist of record titles as if for known-item retrieval or browsing), the investigator retrieved the first and last items listed. In the event of multiple record formats per record, the HTML format was selected.

The resultant 83 records (one agency's GILS featured only 1 record) were printed for ease of study and comparative reference. Their characteristics were assessed and recorded in a relational database for compilation and subsequently transferred to a spreadsheet for analysis using descriptive statistics. A subset of the total was created and subject to identical analysis by filtering the data for values of “US Federal GILS” or “U.S. Federal GILS” in the Controlled Vocabulary-Local Subject Index-Local Subject Term subelement—a state presumed to indicate record-creators’ intention of identifying the record as a “Core record” as delineated in the NARA *Guidelines*. No further operationalization of the “Federal Core” was achieved in this evaluation. The “Core subset” comprised 50% of the total sample.

Table C4-4
Information Object Semantics

Object	Operational Definition	Examples
Administrative Catalog	A locator listing of procedural actions related to the conduct of agency business	FERC's "Directory Of External Information Collection Requirements" PBGC's "Log Of Benefit Termination Plans" USPS's "Index Of Final Opinions And Orders"
Agency Homepage	Information mounted on an HTTP server	"Superintendent of Documents Home Page on the World Wide Web"
Bibliographic Database	An automated information system comprising metadata about bibliographic entities/publications	DOE's "OpenNet" "HUD USER"
Form	A document designed to elicit and transmit specific information from the user to the supplier, respectively	"Request for Registration for Political Risk Insurance" "SSA-1710"
Job Line	A telephonic recording of employment opportunities	"DOI Employment Center"
Miscellaneous Documents In <i>Ad Hoc</i> Collection	Plurality of documents grouped by function or subject	bulletins and memoranda press releases public comments under-described "technical documents" and "reports" update notices letters speeches records
Organization	A set of human resources defined by an agency to provide specific products or services	information center/library research consortium NASA's "Flight Dynamics Facility"
Program	A prescribed set of activities and functions performed to accomplish an objective	report management records management
Publication	Discrete monographic document published one-time or in serial mode to disseminate information	annual report user's manual "The Federal Register" Regulations CD-ROM fact-sheet series procedures manual
Publications Catalog	A fixed, flat (non-machine-searchable) listing of selected or all agency publications	FEMA's "Publications Catalog"
Subject Matter Database	Single, stand-alone automated information system comprising data, records, or multiple documents on technical or administrative subject(s) and/or definable reference themes	Privacy-Act records health risks aviation accidents red cockaded woodpecker
System Of Systems	Macro-AIS comprising or integrating multiple databases and/or single-AISs	DOD's "Enterprise Information System" EPA's "Information Systems Inventory"

BEST COPY AVAILABLE

6.0. METHOD LIMITATIONS AND RECOMMENDATIONS TO FUTURE RESEARCHERS

The primary limitation to the procedures described for analyzing GILS record content is generalizability—the extent to which results can be assumed valid for the entire population of GILS records. The sample was small, less than 2% of the estimated total of approximately 5,000, and the sampling technique was largely convenience-driven due to time constraints. In addition, the method as employed did not provide data concerning differences in record quality among or within agencies' GILS, which might prove useful in estimating the scope of effort required in modifying elements or standardizing the characteristics of element values.

The record content analysis was extremely time-consuming, both in terms of defining mutually exclusive codes for content description and data collection. As noted above, even this small sample involved recognition of presence or absence of thousands of instances of metadata elements as well as examination and description of their values. Much of the labor burden of the current procedure could be alleviated by machine processing—e.g., for element counts, incidence of hypertext, etc. In addition, it is anticipated that the exploratory method described herein will be refined and adapted during subsequent applications, both for assessing the responsiveness of government-wide quality standards for GILS (*vis a vis* the NARA *Guidelines*) and, at the agency level, the quality of GILS record collections.

7.0. CONCLUSION

In summary, the method employed to analyze the content of GILS records proved highly satisfactory in rendering the type of results that would inform the overall evaluation. By providing a bird's-eye view of the "product on the shelf" at a given point in time, this method allows a comparison of planned vs. actual outcomes for quality. Agencies' continuous analysis and reporting of record content will serve well in complementing evaluations of the effectiveness of the NARA *Guidelines*, implementation maturity, and user satisfaction.

APPENDIX C-5
Scripted Online User Assessment Methodology

APPENDIX C-5

Scripted Online User Assessment Methodology

Table of Contents

1.0.	Introduction	1
2.0.	Method Overview	1
3.0.	Objectives.....	1
4.0.	Context Within the Evaluation	2
5.0.	Data Collection and Analysis	2
6.0.	Description of Data Collection Instrument.....	3
7.0.	Participants	4
8.0.	Session Delivery: Logistics and Equipment	5
9.0.	Method Limitations	5
10.0.	Lessons Learned and Recommendations to Future Researchers.....	6
11.0.	Conclusion.....	6

Table C5-1 Scripted Online User Assessment—Participant Profile 4

Attachment C5-1 Introduction and Orientation to
 Scripted Online User Assessment Session 9

APPENDIX C-5 SCRIPTED ONLINE USER ASSESSMENT METHODOLOGY

1.0. INTRODUCTION

The methods used in this research activity served four purposes. First, the online sessions permitted in-process, "front-line," collection of data concerning user assessments of GILS—as opposed to "recollection" of assessments after GILS use. Second, the method was geared toward eliciting highly qualitative responses to a *concept* (i.e., rather than the more traditional aims of user assessments such as quantification of relevant "hits" or usage patterns). Third, the assessment questions attempted to gain insight into cognitive processes of users in the online, networked environment. Last, investigators documented this new exploratory technique and instrument with an expectation that future research will build upon and improve them. Investigators consider these objectives paramount in understanding user perceptions, expectations, and behavior during networked information discovery and retrieval (NIDR), and in advancing the quality of GILS accordingly.

2.0. METHOD OVERVIEW

Graduate and undergraduate student "users" unfamiliar with GILS were oriented to the nature and purpose of their participation by means of a 5-minute verbal introduction by the investigators (see Attachment 1). They were subsequently asked to record answers to more than 50 multiple-choice, free-form expression, and true/false questions as they navigated "real life," "real time" Government Printing Office (GPO) GPO Access and Environmental Protection Agency (EPA) GILS systems according to a scripted set of encounters. The script was based on results of the record content analysis and investigators' ongoing search/retrieval experience with various GILS. The questions were designed to elicit user feedback concerning GILS content and service expectations, record design, orientation in information space, adaptation to the metadata construct (e.g., searching reflexes), and, perhaps most importantly, users' assumptions about GILS—all on the basis of this 1-hour first-exposure to scripted transactions. In addition, investigators conducted debriefing sessions where users were informed generally of GILS scope and purpose and asked to elaborate on intellectual and emotional impressions created by the scripted. The qualitative data from the sessions were entered into a database to facilitate disclosure of patterns related to users' reactions to GILS as a service concept and to GILS product. As with the record content analysis (see Appendix C-5 Record Content Analysis Methodology), investigators recorded suggested improvements to the development and execution techniques for scripted online-user assessment in order to optimize recommendations to agencies interested in adopting the techniques.

The following paragraphs present information concerning the scripted online user assessment research objectives, its context within the overall evaluation framework, data collection/analysis, data collection instrument (script), participants, session delivery, method limitations, lessons learned, and recommendations.

3.0. OBJECTIVES

The scripted online user assessments were conducted to capture users':

- Subjective appraisal of GILS efficacy as a tool for NIDR, including
 - Appreciation of/adaptation to the construct of metadata in the virtual environment
 - Record display/presentation
- Emotional and intellectual reactions to GILS products and services, including
 - Confidence in using GILS
 - Expectations for content and service quality
 - Perceptions of GILS as an information space
- Assumptions about GILS based on a limited first-exposure, including responses to

- Objectives
- Architecture
- Coverage
- Potential.

These objectives were realized in the data collection and analysis activities outlined below.

4.0. CONTEXT WITHIN EVALUATION FRAMEWORK

Both the development of the user assessment script (see Appendix D-5 Scripted Online User Assessment Instrument) and lines of inquiry during the debriefing session were informed by results of previous data collection activities. The content analysis of GILS records (see Appendix E-2 Record Content Analysis Findings, Discussion, and Recommendations) proved particularly useful in identifying the type of GILS encounters most likely to provoke user response. In addition, the content analysis provided a basis for representing the content and quality variations in GILS at macro- and micro levels (i.e., as a government-wide collection and as individual records).

The session provided a degree of triangulation against data collected by way of focus groups and the conference survey. "Users" in the focus groups fell into broad stakeholder categories of agency staff (e.g., records managers and public information officers), public interest groups (perhaps focused on access policy and scope of coverage as much if not more than actual service satisfaction), and professional intermediaries (e.g., librarians). The "users" in the scripted online assessment, however, could be considered "closer to the street" or more representative of "public access"/average citizen end-use requirements.

In addition, as discussed in Section 10 Lessons Learned and Recommendations, the methodology is presented as one in a complementary suite of activities offering benefit in the areas of system requirements definition and quality assurance.

5.0. DATA COLLECTION AND ANALYSIS

The objectives described in Section 3 translated into the following categories of data collection and analysis:

- Content expectations
 - Full-text of documents vs. metadata
 - Subject matter and resource types (objects represented)
 - Information quality (e.g., currency and completeness/level of detail)
 - Scope and extent of collection
 - Record and resource aggregation ("distance" from satisfaction of an information need)
- Service expectations
 - Predictability of results
 - Effectiveness of fielded searching
 - Cause of search logic and "service" errors
 - System/service response time (transaction and record download time)
 - "Comfortability" and satisfaction with GILS purpose and its translation into service
 - Quantity and nature of hyperlinks
 - Implementation policies to yield consistency across GILS systems and other objectives
- Satisfaction with record characteristics
 - Cosmetic appearance
 - Ergonomics (e.g., element display order)
 - Length
 - File formats (ASCII vs. HTML)
- Perceptions of information space and ownership

- Navigation
- Centralization (loci of services and products)
- Implementation architecture (network distribution)
- Availability of resources
- Authority (hierarchy of service and product responsibility for quality)
- GILS nomenclature (use of bibliographical and NIDR terminology in element names and definitions)
- Searching reflexes and relevance judgments
 - Preferences for full-record vs. fielded searching
 - Relevance improvement tactics
 - Requirements for user "sophistication" *vis-a-vis* education and training

These categories, while not of equal weight or under equal levels of control by GILS implementors, compose a backdrop useful for evaluating GILS from the perspective of a capable but "GILS-unaware" online searcher or browser.

6.0. DESCRIPTION OF DATA COLLECTION INSTRUMENT (SCRIPT)

Appendix D-5 Scripted Online User Assessment Instrument presents the instrument used during this research activity. The first 10 items plus 3 on the last page captured demographic and other information about the participants—such as status (e.g., student, private citizen, etc.) and government-information search frequency, methods, resource types, and knowledge/attitudes about government information—provided a context for evaluating expectations and responses.

The content and organization of the script were planned to provide these first-time users with sufficient grounding to execute searching and the widest possible range of GILS experience within a prescribed 1-hour timeframe—a very difficult undertaking—especially given our methodological objective of minimizing effects of user interface and search engine functionality. Users were not permitted to hyperlink to nonscripted help information of any kind nor allowed to repeat searches in order to understand or improve results. (The reasonableness of the former constraint was confirmed during pretest of the user profile sheet below, which indicated that student users are unlikely to consult online help but rather rely on trial-and-error in searching.)

The assessment script directed users to encounter a minimum of two GILS implementations—GPO Access and EPA—and up to ten GILS sites. (Users were provided approximately 10 minutes of nondirected exploration of FedWorld, NTSB, USDA, DOL, GSA, and SBA sites.) Within the two mandatory sites, the script called for six searches (executions of both self-directed and designed query) and subsequent examination of results pages, and examination of four individual records. GPO Access and EPA were selected from among the field of hosts because they afforded an "integrated" vs. "standalone" perspectives (i.e., cross-agency vs. agency-centric models), they are relatively mature implementations, and because users were likely to be familiar with these agencies' missions. The other "optional" sites listed above were selected on the basis of their wide variety of mission, implementation approach (i.e., WAIS vs. HTTP), and types of resources described (e.g., technical reports, databases, consumer information, regulation digests, etc.).

The script attempted to provide users a "fair" or reasonable simulation of personal nondirected use while necessarily constraining movement both to control outcomes, increase cross-user data validity, and save time. With this in mind, the following text broadly outlines the script activities and encounter types in script order:

GPO Access

- Placement in GPO Access GILS space
- Exposure to GILS nomenclature
- Selection of metadata elements (search fields) and formulation of query for a scripted and broad information need
- Assessment of and reaction to search result
- Scripted retrieval of a particular record for comparison with another describing a like resource

- Scripted retrieval of two records having the same title
- Selection of metadata element (search fields), selection of GILS database, and formulation of query for a scripted and very narrow/specific information need (known-item search)
- Assessment of and reaction to search result
- Scripted retrieval of two records having the same control identifier and titles in different languages

EPA

- Placement in EPA GILS space and perceptions relative to GPO Access GILS space/function
- Enumeration of expected EPA GILS content
- Exposure to GILS nomenclature
- Scripted retrieval and subsequent assessment of a particular record's descriptiveness
- Enumeration of assumptions related to authority, extent of coverage, and availability

In addition to these active encounters, the script included a question about users' assumption of the GILS universe extent and 14 post-searching Likert scale (5 points: strongly agree↔strongly disagree, no opinion) questions about the desirability of centralization and standardization, and GILS efficacy in supporting NIDR relative to other models.

7.0. PARTICIPANTS

Participants in the scripted online user assessments were selected from the University of North Texas and Syracuse University student bodies. Invitations and recruitment statements specified prerequisites of familiarity with the Web, Netscape, and fielded searching of online databases. The Texas group comprised four students; the Syracuse sessions involved six users. Table C5-1 summarizes the study group profile, characterized by means of the first eight questions on the script.

Table C5-1
Scripted Online User Assessment—Participant Profile

Background
1 "private citizen", 1 art undergraduate student, 1 political science undergraduate student, 1 history undergraduate student, and 6 library science graduate students
Average of more than 2 years' Internet usage
Government Information Experience
Print sources of government information, on average, searched monthly or less frequently
Frequency of searching online sources of government information varies from weekly to "as required by class"
Reports on government activity/public notices and legislation most frequently sought information
Only one participant had read, heard about, or used GILS (one encounter implied)
Most knew that Federal agencies have libraries
Half of group unaware of the function or purpose of many Federal agencies
Strong agreement that public electronic access to government information is important
Searching Behavior
Self-teaching through trial-and-error predominant method of acquiring/refining online searching skills
Browsing websites or bookshelves more common than use of online help, card catalogs, or application of professional training
One user claims 40% "success" in locating government information by starting with agency homepages

8.0. SESSION DELIVERY: LOGISTICS AND EQUIPMENT

The two hosting universities employed different logistics for staging the script-execution and subsequent debriefing events. The University of North Texas online session was conducted in a graduate computer laboratory reserved for this special use; all users executed the script simultaneously; and all users met immediately upon completion of the online session for a 20-minute, relatively nondirective debriefing about GILS that also solicited additional feedback. At Syracuse, individual users executed the script serially over the course of 3 days at a dedicated workstation within an office and convened for a relatively traditional focus group of more than 1 hour duration early the following week. These variations in approach were not felt to degrade the quality of data but rather enhanced investigators' appreciation of the methods' flexibility and potential site-specific adaptation.

Baseline resources to conduct the sessions included workstations with:

- Pentium processor
- Netscape 2.0 or higher:
 - Options/Security Preferences set to disable security popups
 - Bookmarks set to script sites, or session homepage with links to script sites
- Well-behaved mouse
- Mousepad
- Clean screen
- Adequate seating
- Adequate lighting
- Adequate work area (recommended print orientation for script is landscape not portrait).

Other required materials included:

- (2) No. 2 sharp pencils with erasers
- Human subjects consent forms.

Prior to beginning the online session, users were asked to complete the demographic portion of the script and listen to a brief overview of session activities and expectations, including admonishments against deviating from the printed instructions and reassurance that research assistants were available to deal with the inevitable vagaries of the Internet such as error messages, site inaccessibility, etc. and to clarify script instructions.

9.0. METHOD LIMITATIONS

The exploratory technique of real-time scripted GILS user assessments presents several challenges to investigators and participants. First and foremost is development of a script that sieves out extraneous variables such as user interface and search engine functionality. Second, the script must be understandable by users having a wide variety of learning styles and reading comprehension levels. Third, the instrument must facilitate predictable and common searching "paces" among users with varying types and speeds of searching reflex. Last, the script must contain a narrow margin for both user- and system-induced errors.

In addition, participants must:

- be familiar with the Web and selected browser
- have at least rudimentary online search skills
- be capable of following instructions independently
- be capable of recording thoughts-in-process
- be willing to contribute up 2 hours' participation (if debriefing/focus group is included).

The method also demands a relatively high resource commitment in terms of standardly equipped computing workstations and in terms of human skill to create scripts, to coordinate session delivery, to lead and record focus groups, to analyze largely qualitative data, and to integrate those data into the larger picture of research objectives.

For users, the requisite constraint against "surfing" around the information space of GILS—as both a dedicated (self-contained) system and as an intellectual and/or virtual link to other resources—is unnatural and frustrating, and may lead to a perception that the speed and accuracy of mouseclicking is a hidden agenda of the research. It is important to explain to users that the tight scripting is critical to internal validity in terms of "investigators must be certain that they are assessing *specifically* and *only* what they set out to assess."

10.0. LESSONS LEARNED AND RECOMMENDATIONS TO FUTURE RESEARCHERS

During the development, delivery, and assessment of the scripted user online-assessment method, investigators identified the following areas for improvement.

Script Development

Investigators found that a deliberate review of GILS record content (such as that performed in this evaluation study) not only informs the specification of evaluative criteria and selection of supporting scripted encounters but helps focus the script creator on *service output* characteristics rather than system-dependent variables such as interface and search engine performance. In addition, pretesting of the script as close to actual delivery time as possible is essential to clarify wording, optimize organization of the material, time the session, and confirm hyperlinks.

Facility Readiness

Every attempt should be made to ensure the availability and adequacy of required resources (specified above) in order to maximize online time. The "required resource" of network response time, however, proved difficult to manage in the current study. (University of North Texas delivered a 6:00 p.m. session that featured instances of poor response comparable to that at Syracuse sessions run at 11:30 a.m.). Control or isolation of this variable is desirable to increase the reliability or internal validity of the script (i.e., to ensure that the user is responding to targeted GILS features); however, a naturalistic study of GILS should account for the effects of response time on user satisfaction.

Supporting/Validating Activities

The results of a session such as those described above are obviously limited in generalizability—a valid cross-sampling of potential GILS user populations is most likely impossible, the script as instrument will inevitably carry the stamp of its creator(s) perception of GILS, and time does not permit adequate comparison by users of the multiple GILS approaches and output. Given these inherent limitations to the method, investigators strongly recommend that online assessments be supplemented (or "triangulated") with other user-oriented research methods such as a focus group, a library "site visit" for unobtrusive observation, "talk-aloud protocol" free-form search sessions, online user-satisfaction surveys, panel studies, and personal interviews.

11.0. CONCLUSION

The use of a scripted online user assessment of GILS proved complementary to the study's other evaluation methods and provided invaluable data about users' first impressions of current agency implementations. Advantages of the technique include recency of data, its applicability to a potentially diverse base of GILS users—from librarians and other intermediaries, to targeted searchers (e.g., students completing assignments), browsers of information (e.g., subject-oriented novices), and public-access advocates concerned with accountability and scope of coverage.

In addition, user-assessment scripts may be tailored specifically to a number of GILS implementation characteristics—e.g., current and planned search features, record presentation specifications, and depth/breadth of resource coverage—during critical stages in service maturity.

Data generated from actual users, especially real-time data, are essential in creating user-responsive, realistic objectives for GILS and in evaluating the service's performance to those objectives. It is hoped by the investigators that development and deployment of user-satisfaction instruments, such as this scripted assessment, will be recognized by agencies as worthy of the considerable investment required, and that methodologies will be shared in the spirit of continuous improvement.

Attachment C5-1**INTRODUCTION AND ORIENTATION TO
ONLINE SCRIPTED USER ASSESSMENT SESSION**

Introduce oneself and other research team member(s) if present.

Paraphrase the following information when addressing study participants:

This session will last about 1 hour and 15 minutes and be divided into roughly three parts. First, we will ask you to fill out a brief questionnaire about your online searching background. Then, for most of the session, you will use a set of written instructions to navigate through a Federal government information locator service. We'll talk about that more in a moment. Last, we will conduct a 15-minute debriefing in the conference room across the hall. This will give us a chance to listen to your impressions about the search session and the locator service in a relaxed, informal setting. If you haven't yet signed the consent form, please do so now.

The written instructions you will follow to assess the information system are designed to draw out your opinions, impressions, and reactions—*not to test your knowledge about the government or about the World Wide Web or about database design*. With this in mind, please perform each exercise at a determined and positive rate rather than leisurely rate that you might normally use when searching on the web. We realize that this guideline imposes some unnatural restrictions on your curiosity, but after you perform the written instructions you will have 5 or 10 minutes to explore the system in a free-form fashion. Please take each question at face value and record your reactions honestly. Don't concern yourself with complete sentences or grammar, and feel free to use common abbreviations. Don't censor yourself! Again, we need your initial responses from the point of view of a consumer of a service.

We emphasize that it is essential that you follow the instructions exactly, clicking on icons and hypertext only when asked. This not only saves time but ensures that each of you is experiencing the same feedback from the system. Also, do not look ahead at questions or go back to change completed answers; take one at a time. However, if at any time you find the instructions unclear or incorrect, or if the information on your screen seems out of sync with the instructions, please raise your hand right away, and one of us will get things back on track. We will occasionally look over your shoulder to see where you are in the instructions and may ask that you try to answer the questions more quickly.

At about the midpoint of the written instructions, you will find our request that you take a short break. Please use this time to clear your mind of the system for a few minutes by visiting the restroom or water fountain directly across the hall or standing up to stretch and refocus your eyes. We don't want anyone to experience information overload!

When you have answered all the questions in your instruction booklet, please hand it to one of us and head toward the snacks waiting in the conference room around the right-hand corner from the restrooms.

Relax and have fun with this—take this opportunity to tell system designers what you really think of a service! Your responses are completely confidential. We feel very fortunate to have willing and interested volunteers, and *thank you* for taking this time to share your opinions with us.

Note to researcher: Script Step 7 calls for your assistance if user retrieves "Helpful Hints" as one of three hits. Direct him/her to re-do the search with the term (with quotes) "native american" OR indian.

APPENDIX C-6
Web Server Transaction Log Analysis Methodology

APPENDIX C-6

Web Server Transaction Log Analysis Methodology

Table of Contents

1.0. Introduction	1
2.0. Web Server Transaction Log Analysis as a Research Technique in Support of the Project Goals	1
3.0. Description of Web Server Transaction Log Analysis	1
4.0. Transaction Log Analysis Data Collection and Data Analysis Activities.....	2
5.0. Limitations of Transaction Log Analysis as a Research Method.....	2
6.0. Privacy Issues and Conclusions.....	3

APPENDIX C-6 WEB SERVER TRANSACTION LOG ANALYSIS METHODOLOGY

1.0. INTRODUCTION

The assessment and evaluation of electronic networks and network-based resources is increasing in scope and application (Bertot & McClure, 1996a, 1996b; McClure & Lopata, 1996). Web server transaction log file analysis is a network-based assessment technique that is particularly useful when performed in conjunction with other ongoing evaluation activities. The investigators designed an experimental data collection technique to analyze usage of GILS records on a Web server at the Environmental Protection Agency (EPA).

The intent of the Web server log analysis involved four purposes. One intent was to determine the overall Web site traffic including the location of users, the portions of the site accessed, and the number of document downloads. The second purpose was to determine the use of the Web site GILS directory traffic including the location of users, portions of the site accessed, and number of document downloads (both hits and accesses). The third purpose was to experiment with developing new log analysis techniques that go beyond domain, hit, and browser counts. The fourth purpose was to assist Federal agencies that operate Web-based GILS servers to develop, implement, and maintain ongoing log file analysis.

2.0. WEB SERVER TRANSACTION LOG ANALYSIS AS A RESEARCH TECHNIQUE IN SUPPORT OF THE PROJECT GOALS

Federal agencies make increasing use of the Web to provide access to Federal government information sources and in particular, to provide access to GILS records. In supporting GILS on the Web, Federal agencies have several important concerns that the investigators explored. These considerations include knowing what a server's traffic load is and the agency's overall ability to meet the demands of that traffic; knowing what a particular server's user community includes (e.g., accessing host IP address, browser, and operating system); knowing what users do while using the server; knowing both at what point and from where users accessed and left the server; and finally, knowing what problems users encountered during their server sessions. The investigators sought to develop a method by which Federal agencies could measure these indications of use to better manage their resources.

The investigators found that available log analysis software packages, commercially on the market, are generally inadequate to analyze log files in a variety of ways. The investigators reviewed multiple Web analysis software packages and analyzed them against four criteria: the ability to provide global and directory specific Web server analysis; the ability to distinguish between hits and accesses; the ability to determine user-specific actions though Web site session, and the ability to distinguish between unique and total referrals. None of the packages reviewed met all four criteria.

The investigators developed PERL-based scripts to analyze EPA log files that would provide all the required analysis capabilities. With these newly developed tools, the investigators created a mechanism that agencies could use to determine whether Web-based services are meeting the intended mission of the agency to provide public access to government information. By analyzing logs of user transactions, the investigators also attempted to asses a measure of user needs.

3.0. DESCRIPTION OF WEB SERVER TRANSACTION LOG ANALYSIS

The technique of Web server log analysis involved a three-fold process that included determining the types of information server administrators and decision makers need; developing a program that can parse through, manipulate, and present value-added information from the log files; and analyzing the information generated from the program. The investigators used four different log files which are automatically generated by Web servers (Rubin, 1996; Noonan, 1996; Novak and Hoffman, 1996). These four files are the access logs (e.g., hits), agent log

(e.g., browser, operating system), error log (e.g., download aborts), and referer logs (e.g., referring links). These log files are text files that can range in size from 1KB to 100MB, depending on the traffic at a particular site.

Critical to understanding the type of data contained in these files is the distinction between a hit and an access. A hit is any file from a web site that a user downloads. If a user downloads a web page that has 6 images on it, then that user "hit" the web site seven times (6 images + 1 text page). An access (sometimes called a page hit) is an entire page downloaded by a user regardless of the number of images, sounds, or movies on the page. If a user downloads a web page that had 6 images on it, then that user just accessed one page of the web site.

One failing in most of the commercially available log analysis software packages is that the software counts the number of hits a server receives, rather than the number of accesses. The hit count reflects the number of items (e.g., images) downloaded when a user accesses a particular page. If a site has an image file, such as a Federal agency logo on multiple pages, that image will more than likely be the most frequently downloaded "hit" item on the site. Analysis information such as this is relatively useless in determining the site's actual usage.

4.0. TRANSACTION LOG ANALYSIS DATA COLLECTION AND DATA ANALYSIS ACTIVITIES

The investigators selected the EPA Web server from which to collect transaction log analysis files. Both the Department of Defense (DoD) and EPA offered to work with the investigators and the choice of agency became a decision of convenience and size. The size of the files at EPA ranged from 8 megabytes to 26 megabytes each on a daily basis, and this size was significantly smaller than the file sizes from DoD. In all, investigators used approximately 560 megabytes of log file data.

The investigators analyzed the access log, the agent log, the error log, and the referrer log. Log files were collected on February 8, 1997 and February 15, 1997. Each log file included a week's worth of transactions. The resulting output (Web log file analysis PERL scripts and log files) together consumed approximately 1 gigabit of storage.

The development and pretesting of the PERL scripts required considerable effort. The Syracuse University script development team required the equivalent of 240 man-hours developing the scripts. An additional 100 man-hours were required to pre-test the scripts using several different log files from different servers, including a test data set from the Federal agency GILS server. Running the scripts on the 14 day period of EPA log files and outputting the analysis into a usable format required an additional 100 man-hours. In total, therefore, the PERL script development process consumed approximately 420 man-hours.

The analysis of the EPA log files was performed on a Pentium 150 MHZ computer with 32 MB of RAM, and the analysis of each of the four daily log files took approximately 40 minutes.

5.0. LIMITATIONS OF TRANSACTION LOG ANALYSIS AS A RESEARCH METHOD

This method is exploratory and as such, is subject to further development. The use of this method encountered a few limitations.

The investigators had no guarantee that the files they received were complete data sets. There is a need to post the file size of the log files directly from the server such that agency staff responsible for this analysis can verify the file size against the downloaded files.

The investigators stored two weeks of log files from EPA as well as the PERL scripts. The resulting files took up nearly a gigabit of hard drive space. If an agency were to maintain this type of analysis on an ongoing basis, there is a need to dedicate a machine with adequate hard disk space for the task. It is also necessary to have a backup server or tape backup of the script and the log data files.

The investigators underestimated the number of daily referrals that the EPA server received. From analyzing the log analysis data, it is clear that a number of referrals came from search engines. The PERL scripts were not written to

extract this information. Future development of the scripts can help to determine not only what percentage of referrals come from search engines, but what search engines users tend to use and what search terms users enter.

In a more general context, it is important to interpret and consider log files as one component of a larger assessment activity for network services. While log files can provide Web administrators and others with critical server-related data, log files do not reflect user-based impact and outcome measures. As such, log files combine both user and technical perspectives on Web services.

The distinction between "hits" (downloads on an html page) and accesses (a downloaded html page) is critical. Software that counts only "hits" will not reflect the true nature of the site's use.

Agency use of commercially available transaction log analysis software may not readily support this distinction. Web administrators should not retrofit their log file analysis to the capabilities of this type of software.

Gaining access to and analyzing Web sever log files requires planning, coordination, and accountability. To engage in log file analysis activities, there needs to be a delegation of responsibility for making the files available (onsite or remotely), performing the analysis (onsite or remotely), interpreting the analyzed data, and reporting the findings. Moreover, such analysis needs to be performed and reported on an ongoing and regular basis.

There is a need to resolve these issues and move the ability to perform log file analysis forward. Log file data can provide user-based measures of Web-based resources if performed on a regular basis, incorporated into other electronic network assessment activities, and interpreted correctly.

6.0. PRIVACY ISSUES AND CONCLUSIONS

A major issue connected with data collection of HTTP transaction log analysis is privacy. In some cases, it is possible to trace directly back to a user, depending on the method of access a user has to a Web site. Web service providers need to develop policies as to how such data are to be used, if at all. This issue is particularly troublesome for public sector organizations, as such capabilities may violate privacy laws.

Web server transaction log analysis was an experimental data collect technique used by the investigators. This means of collecting data is relatively new and is one which will grow in importance to Federal agencies.

**APPENDIX D-1
Site Visit Profile Sheet
(Example)**

GILS Evaluation Project Site Visit Questionnaire

1. Current Position -- Title and Description: _____
2. How long have you been in your present position? _____ years
3. Which one category *best* describes the area in which you work?

<input type="checkbox"/> Chief Information Officer <input type="checkbox"/> Program Office/Project Management <input type="checkbox"/> Procurement/Contracting <input type="checkbox"/> Legal/Legislative <input type="checkbox"/> Other Information Resources Management – please describe: _____ <input type="checkbox"/> Contractor -- please describe services provided: _____ <input type="checkbox"/> Other -- please describe: _____	<input type="checkbox"/> Computer Systems <input type="checkbox"/> Records Management <input type="checkbox"/> Library/Information Center <input type="checkbox"/> Public Information Office
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
4. Please list several terms that best describe your area(s) of professional expertise: _____
5. How familiar are you with the following (circle the appropriate number):

	Very	Familiar	Not Familiar		
a. Federal GILS policies	1	2	3	4	5
b. Your agency's GILS policies	1	2	3	4	5
c. NARA's guidelines for GILS records creation	1	2	3	4	5
d. OMB Bulletin 95-01 on GILS	1	2	3	4	5
e. Z39.50 standard	1	2	3	4	5
f. FIPS Publication No. 192	1	2	3	4	5
g. Paperwork Reduction Act of 1995 section establishing GILS	1	2	3	4	5
- 6a. Please identify your favorite online (WWW or otherwise) *alternative* to GILS when trying to locate government information: _____
- 6b. Why do you especially like this *particular* source? _____
7. What do you think are the two primary barriers issues facing agencies in planning and implementing GILS?
 - a) _____
 - b) _____
8. What do you think are the two most important lessons you learned so far in your GILS implementation?
 - a) _____
 - b) _____
9. What are the two most important decisions facing your agency in future GILS activities and operations?
 - a) _____
 - b) _____
- 10a. Did you receive training in any aspects of GILS? _____ Yes _____ No
- 10b. Describe briefly the training you did receive: _____

**APPENDIX D-2
Focus Group Profile Sheet
(Example)**

GILS Evaluation Project Focus Group Questionnaire

1. Current Position -- Title and Description: _____

2. How long have you been in your present position? _____ years

3. Which one category *best* describes the area in which you work?

<input type="checkbox"/> Chief Information Officer <input type="checkbox"/> Program Office/Project Management <input type="checkbox"/> Procurement/Contracting <input type="checkbox"/> Legal/Legislative <input type="checkbox"/> Archives <input type="checkbox"/> Hardware/Software Developer/Vendor <input type="checkbox"/> Other Information Resources Management -- please describe: _____ <input type="checkbox"/> Contractor -- please describe services provided: _____ <input type="checkbox"/> Other -- please describe: _____	<input type="checkbox"/> Computer Systems <input type="checkbox"/> Records Management <input type="checkbox"/> Library/Information Center <input type="checkbox"/> Public Information Office <input type="checkbox"/> Public Interest Group <input type="checkbox"/> Library Organization
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

4. Please list several terms that best describe your area(s) of professional expertise: _____

5. How familiar are you with the following (circle the appropriate number):

	Very	Familiar	Not Familiar		
a. Federal GILS policies	1	2	3	4	5
b. Your agency's GILS policies	1	2	3	4	5
c. NARA's guidelines for GILS records creation	1	2	3	4	5
d. OMB Bulletin 95-01 on GILS	1	2	3	4	5
e. Z39.50 standard	1	2	3	4	5
f. FIPS Publication No. 192	1	2	3	4	5
g. Paperwork Reduction Act of 1995 section establishing GILS	1	2	3	4	5

- 6a. Please identify your favorite online (WWW or otherwise) *alternative* to GILS when trying to locate government information: _____

- 6b. Why do you especially like this *particular* source? _____

7. What are the two most important benefits that the Federal government should achieve by implementing GILS?
 - a) _____
 - b) _____

8. What are the two greatest barriers or threats to the success of GILS?
 - a) _____
 - b) _____

9. What are the two most important lessons learned from GILS as an example of Federal information policy?
 - a) _____
 - b) _____

- 10a. Did you receive training in any aspects of GILS? _____ Yes _____ No
- 10b. Describe briefly the training you did receive: _____

**APPENDIX D-3
GILS Conference Survey Instrument**

**ASSESSMENT OF KEY ISSUES RELATED TO THE
GOVERNMENT INFORMATION LOCATOR SERVICE (GILS)**

Instructions: Please provide the information requested below and return the questionnaire to the boxes placed at the back of this room.

Part I: Background Information

1. What is your definition of GILS -- Government Information Locator Service?

2. Which of the following best describes the setting in which you work?

() Federal agency	() Private for-profit organization
() Private non-profit organization	() Other (please describe)
() State or local government	

3. _____ years of experience in this setting.

4. Which one category best describes the field in which you work?

() Chief Information Officer	() Library/Information Center
() Research and Development	() Information Resources Management
() Software Developer	() Other (please describe)
() Sales/Marketing	

5. Which one category best describes your involvement with or interest in GILS?

() Implementor of GILS	() GILS record creator
() User of GILS	() Policymaker
() Technical standards developer	() Other (please describe)
() Information reseller	

6. For government agency employees ONLY:

a. Our agency operates a GILS server
() yes () no () don't know

b. Our agency submits GILS records to another agency for listing
() yes () no () don't know

c. My unit's GILS is part of a larger department/bureau/agency GILS effort
() yes () no () don't know

7. How familiar are you with the following (circle the appropriate number):

	Familiar				Not Familiar
a. Federal GILS policies	1	2	3	4	5
b. Your agency's GILS policies (if government employee)	1	2	3	4	5
c. NARA's guidelines for GILS records creation	1	2	3	4	5
d. OMB Bulletin 95-01 on GILS	1	2	3	4	5
e. Z39.50 standard	1	2	3	4	5
f. FIPS Publication No. 192	1	2	3	4	5
g. Paperwork Reduction Act of 1995, section establishing GILS	1	2	3	4	5
h. The World Wide Web	1	2	3	4	5

8a. Please identify your favorite online alternative (Web or other) to GILS when trying to locate government information:

8b. Why do you especially like this particular source?

Part II: Key Issues

Please circle the number or letter that corresponds with the degree to which you agree or disagree:

	Strongly Agree		Strongly Disagree	Don't Know
9. I am able to describe GILS accurately and fully to others.	1	2	3	4 5 X
10. A purpose of GILS is to improve public access to government information.	1	2	3	4 5 X
11. There is adequate policy guidance to direct the development and operation of GILS.	1	2	3	4 5 X
12. There is adequate technical implementation guidance for developing and operating GILS.	1	2	3	4 5 X
13. GILS records represent the complete information resources of an agency.	1	2	3	4 5 X
14. More work needs to be done on the technical standards to support GILS operations.	1	2	3	4 5 X
15. GILS is so valuable that, if it didn't exist, we would have to create it.	1	2	3	4 5 X
16. Every agency Web homepage should have a link to the agency's GILS.	1	2	3	4 5 X
17. GILS has improved public access to Federal information.	1	2	3	4 5 X
18. A purpose of GILS is to help agency officials better manage agency information.	1	2	3	4 5 X
19. The World Wide Web reduces the need for GILS.	1	2	3	4 5 X
20. The appropriate technology is readily available for implementing GILS.	1	2	3	4 5 X
21. GILS should be expanded throughout the Federal government.	1	2	3	4 5 X

Please circle true or false for the following:

22. I search GILS several times per day in my everyday work.	True	False
23. I often refer people to GILS when providing information.	True	False
24. I find useful information when I use GILS.	True	False
25. I often find links to GILS on the Web.	True	False

Part III: Additional Comments

26. Please use the space below to offer any additional comments, pose questions, identify key issues, or discuss topics not mentioned here that the study team should explore:

**APPENDIX D-4
Record Content Analysis Instrument
(Database Fields)**

APPENDIX D-4
Record Content Analysis Instrument
(Database Fields)

The following table represents the Microsoft Access© database structure devised and used to collect data for the record content analysis. The first column presents the data category, or criteria assessed; the second column shows the range of allowable data entries; and the third column contains information concerning operationalization of the criteria to assist in investigators' recall of rationale.

Criteria*	Values**	Data Coding Notes
Server	<ul style="list-style-type: none"> - http - index-based 	
Host	<ul style="list-style-type: none"> - FedWorld - GPO - nonbrokered 	
ASCII text available	y/n	
HTML available	y/n	
PDF available	y/n	
SGML available	y/n	
File format errors present	y/n	word wrap, uncontrolled indents, missing text
Spelling errors/typos present	y/n	
All dates in yyyyymmdd format	y/n	examine all dates in record
Acronyms	<ul style="list-style-type: none"> - defined - not defined - nu 	
Legislative cite utilized	y/n	
Capitalization style	<ul style="list-style-type: none"> - all caps - elements only cap - mixed - other - sentence case 	
Indentation present	y/n	regardless of adherence to <i>Guidelines</i> -specified style
Total number of elements utilized	integer	element is presented and populated (including where "none", "n/a", or similar acknowledgment of an absence of substantive content is present)
Blank elements present		element/subelement label present and no value present
Locally defined elements	actual value	any element/subelement not mentioned in <i>Guidelines</i> ; count null value if element label present
Hotlinks	element label	of value containing hypertext
Elements in preferred display order	y/n	per FIPS 192

(continued)

Criteria*	Values**	Data Coding Notes
Object represented	<ul style="list-style-type: none"> - administrative catalog - agency homepage - bibliographic database - form - job line - miscellaneous documents in ad hoc collection - organization - program - publication - publications catalog - subject matter database - system of systems - unknown 	see semantics in Appendix C-4 Record Content Analysis Methodology
Aggregation	<ul style="list-style-type: none"> - aggregated object represented - discrete object represented - object aggregates metadata - record aggregates objects - unknown 	see semantics in Appendix C-4 Record Content Analysis Methodology; "object aggregates metadata" overrides "discrete object represented"—e.g., discrete publications catalog
Container	<ul style="list-style-type: none"> - broadcast (radio or TV) - CD-ROM - dialup - email - fax - ftp - gopher - listserv - microform - multiple - print - unknown - video - voice/telephone - web 	

(continued)

Criteria*	Values**	Data Coding Notes
Title descriptive	y/n/nu	per <i>Guidelines</i> : "...conveys the most significant aspects of the referenced resource...provide[s] sufficient information to allow users to make an initial decision on likely relevance...convey[s] the most significant information available, including the general topic area, as well as a specific reference to the subject."
Originator utilized	y/n	
Originator subelements utilized	y/n	"yes" even if no subelement label present
Originator subelements labeled	y/n	
Index terms-controlled utilized	y/n	
Thesauri	actual value	
Local subject terms	<ul style="list-style-type: none"> - actual value if "US Federal GILS" or "U.S. Federal GILS" - nu 	
Abstract descriptive	y/n/nu	
Purpose utilized	y/n	
Agency program utilized	y/n	
Availability-distributor utilized	y/n	
Availability-resource description utilized	y/n	
Availability-order process utilized	y/n	
Source of data utilized	y/n	
Methodology utilized	y/n	
Access constraint	<ul style="list-style-type: none"> - "none" - named - nu 	use "named" when value is substantive
Use constraint	<ul style="list-style-type: none"> - "none" - named - nu 	use "named" when value is substantive
Point of contact	<ul style="list-style-type: none"> - job title - nu - office - other - personal name 	
Schedule number	actual value	
Control ID utilized	y/n	
Record source subelements utilized	y/n	"yes" even if no subelement label present
Record source subelements labeled	y/n	
Date of last modification	actual value	

* Criteria are arranged in convenience order for data collection using visual examination of printed record.

** y = yes; n = no; nu = element not utilized.

APPENDIX D-5
Scripted Online User Assessment Instrument

User Profile: Scripted Online Assessment of GILS

1. How do you chiefly acquire or refine your online searching skills? (Circle one.)

 - a. Self-teaching by trial and error
 - b. Professional training
 - c. Reading online Help manuals
 - d. Applying knowledge of database design
 - e. Other (please describe): _____

2.

	Daily	Weekly	Monthly	Other (please specify)
a. How often do you search print sources of government information?				
b. How often do you search online sources of government information?				

- 3a. Have you ever read or heard about the U.S. Government Information Locator Service (GILS), or actually used it? ____ yes ____ no
- 3b. If you answered "yes" in Step 3a, how did you discover GILS?

4. Approximately how long have you used the Internet? ____ months ____ years
5. Please circle the letter that most closely describes your current status:

 - a. Political Science undergraduate student
 - b. Library Science graduate student
 - c. Political Science graduate student
 - d. Private citizen
 - e. Other (please describe): _____

6. What types of government information do you seek most frequently? (Circle up to three):

 - a. Research
 - b. Case law
 - c. Historical
 - d. Statistics
 - e. Regulations
 - f. International relations
 - g. Legislation
 - h. Budget and economic news
 - i. Reports on government activity/Public notices

7. Please circle letter(s) matching your experience seeking government information, in print or online:

- a. I usually need help from a librarian or other intermediary to get started
- b. I browse websites or bookshelves to find information
- c. I nearly always find just what I need
- d. I find user's instructions sorely lacking for most resources
- e. I have a few favorite sources that I have learned to use
- f. I find that government information sources change often
- g. I begin my search with using a card catalog or online index
- h. I avoid searching government sources directly whenever possible, and rely on secondary reports such as newspapers or CNN

8. If you were to enter search terms into an online "information locator service", what would expect in return? (Circle all that apply.)

- a. Full text of documents that contain the information I seek
- b. A "frequently asked question" (FAQ) list with answers
- c. Statements about where relevant information is stored and how to obtain it
- d. Relevant document titles only
- e. Names of experts in the subject
- f. A list of related, controlled subject terms from which to choose
- g. Relevant database names
- h. Abstracts or digests from relevant documents
- i. Other (please describe): _____

Thank you for providing this background information

**An Evaluation of the Federal Government's Implementation of the Government Information Locator Service (GILS)
User Online-Assessment Script**

Please click on Netscape's **Bookmarks** button and select the first item: *Government Information Locator Service*.

Receive *Government Information Locator Service (GILS) Online via GPO Access*.

DO NOT CLICK ON ANY BUTTONS, ICONS, OR HYPERTEXT UNLESS SPECIFICALLY INSTRUCTED.

Please examine this entire page.

Use the scrollbar next to the box labelled **Make your selection(s)**.

1. There are many options listed underneath **Individual Agency GILS databases on GPO Access**. What do you think these might represent? (Circle all that apply.)
 - a. Publishers/distributors of information
 - b. Information creators
 - c. Internet server "mirror" locations
 - d. Other (please describe): _____

Click on the blue hypertext phrase **fielded search**.

Receive *Search GILS Records by Fields Online via GPO Access*.

Please direct your attention to the screen section labelled **Select one or more of the following fields (mandatory GILS core elements) to search**.

2. Would you assume that "mandatory GILS core elements" means that these fields always contain data? _____ yes _____ no
3. Below, please rate how confident you would feel using the options presented under **Select one or more of the following fields "mandatory GILS core elements" to search**: (NOTE: DO NOT CLICK ON BLUE HYPERTEXT FIELD NAME)

231

230

	Certain	Unsure	Unsure
Abstract		Originator	
Access Constraints		Point of Contact for Further Information	
Agency Program		Purpose	
Availability		Record Source	
Control Identifier		Sources of Data	
Controlled Vocabulary		Spatial Reference	
Cross Reference		Schedule Number	
Date of Last Modification		Supplemental Information	
Local Subject Index		Time Period of Content	
Methodology		Title	
Original Control Identifier		Use Constraints	

On your screen, click on the blue hypertext field name you are most "unsure" of from Step 3. (If you did not use "unsure", please click the field name about which you feel least "confident".)

Receive and read the definition of this field.

4. Does this definition affect your confidence in using this field for searching? increases it decreases it no change

232

233

Please click Netscape's Back button.

Receive Search GILS Records by Fields Online via GPO Access.

5. Please check below which ONE field you would select for locating information about native americans.

- Abstract
- Access Constraints
- Agency Program
- Availability
- Control Identifier
- Controlled Vocabulary
- Cross Reference
- Date of Last Modification
- Local Subject Index
- Methodology
- Original Control Identifier
- Originator
- Point of Contact for Further Information
- Purpose
- Record Source
- Sources of Data
- Spatial Reference
- Schedule Number
- Supplemental Information
- Time Period of Content
- Title
- Use Constraints

In the scrollbox labelled **Make your selection(s)**, click on All Records on GPO Access GILS Site (this is the most comprehensive option).

In the section labelled **Select one or more of the following fields (mandatory GILS core elements) to search**, click inside the checkbox for the field you selected in Step 5.

Click Submit.

Receive Search GILS Records by Relevant Term.

Please read the paragraph on your screen that starts with the phrase **ENTER SEARCH TERMS IN THE SPACE ABOVE. DO NOT LINK TO INSTRUCTIONS AND EXAMPLES.**

In the screen textbox, type any search term(s) you think are relevant to the concept of native americans. Please double-check your spelling.

Click Submit.

Receive GILS Search Results.

6. How many "total hits" did you receive? _____

7. What is your reaction to the *number* of "hits"?

- a. Pleasantly surprised
- b. Overwhelmed but willing to examine some of the hits more closely
- c. Overwhelmed but willing to start over with more specific search terms
- d. Not surprised
- e. Disappointed but willing to examine the hits more closely
- f. Disappointed but willing to start over with different search terms
- g. Frustrated--I would abandon use of GILS at this point

If you received a "hit" called Helpful Hints for Searching GILS Records, please raise your hand and a research investigator will provide you with further instructions.

Please use the scrollbar to quickly scan the list.

8. If you were to select one or more hits for closer examination, which factor would mostly likely influence your selection? (Circle one.)

- a. Order of appearance (select first item first)
- b. Appearance of search terms in Title
- c. Size
- d. Score
- e. Format
- f. Other (please specify): _____

For the hit that seems most promising, click on the blue hypertext HTML if present; otherwise click on the blue hypertext text.
SCAN the returned information BRIEFLY. DO NOT CLICK ON ANY HYPERTEXT IN THE RECORD.

9. Please characterize and explain your single FIRST REACTION to this record:

- a. Surprised, because...
- b. Confused, because...
- c. Interested, because...
- d. Disappointed, because...
- e. Frustrated, because...

10. Please describe anything you consider to be peculiar or in error:

Please click on Netscape's **Bookmarks** button and select the first item: *Government Information Locator Service*.
Receive *Government Information Locator Service (GILS) Online via GPO Access*.
In the scrollbox **Make your selection(s)**, click on **International Trade Commission**.
In the textbox **Enter search term(s):**, type the word **library**.
Click on **Submit**.
Receive *GILS Search Results*.
Click on the blue hypertext [HTML](#) for hit [1] **GILS: U.S. International Trade Commission. Library Services.**

Receive and SCAN the GILS record. DO NOT CLICK ON ANY HYPERTEXT IN THE RECORD.

11. What do you think would happen if you were to click on this record's hypertext title? DO NOT TAKE THIS ACTION.

- a. I would jump to the ITC website
- b. I would connect to ITC's online library catalog
- c. A list of ITC library staff contacts would appear
- d. I would link to a fuller/longer version of this record
- e. I would be given a list of library services such as interlibrary loan, photocopying, and research assistance
- f. Other (please specify): _____

12. Please rate your feeling about the *length of this record relative to your satisfaction*:

- a. Too short; it doesn't present enough detail.
- b. Too long; it provided more than I needed to know.
- c. Just right; it presented the necessary information.

	Yes	No
a. Do you think the search term "trade agreements" would have produce a "hit" on this record?		
b. If you were looking for information about the ITC library, would this record satisfy you?		
c. Did you know that Federal agencies have libraries?		
d. Would you have expected to find a Government Information Locator Service record that describes a library?		

Please click on Netscape's Back button twice to return to *Government Information Locator Service (GILS) Online via GPO Access.*

In the scrollbox **Make your selection(s)**, click on **Merit Systems Protection Board**.

In the text box **Enter search term(s):** leave in the word library.

Click on **Submit**.

Receive and scan *GILS Search Results*. DO NOT LINK TO ANY RECORDS.

14. Why do you think two apparently identical results have been returned? (Circle all reasonable possibilities.)

- a. The search term was too broad
- b. The system has made a error in searching or retrieval.
- c. The titles of the two records have been shortened for this display; they are actually different
- d. Both records describe the Merit Systems Protection Board (MSPB) library, but were created by different agencies
- e. The person who created the duplicate record was unaware that a record already existed.
- f. The instructions I followed for this search are incorrect or incomplete.
- g. Other (please describe): _____

Click on the blue hypertext HTML for hit [1] GILS: MSPB LIBRARY.
Receive and SCAN the GILS record.

15. Please describe anything you consider to be peculiar or in error:

16. Of the two agencies' records describing libraries; which is best? _____ MSPB ITC no preference

17. What characteristics distinguish the two records in your mind? (Circle all that apply.)

- a. Accuracy
- b. Format
- c. Completeness
- d. Presence of hotlinks
- e. Currentness
- f. Consistency
- g. Other (please specify): _____

Please take a minute to get up and stretch or to visit the water fountain or restroom.

If required, eliminate the screen saver graphics by clicking anywhere on your screen.

Please click on Netscape's **Bookmarks** button and select the first item: *Government Information Locator Service*.

18. Below, describe how you would use GILS to find the TOLL-FREE NUMBER FOR ORDERING A STATEMENT OF EARNED SOCIAL SECURITY BENEFITS

- a. What would you choose in the **Make your selection(s)** scrollbox? _____
- b. Would you use "fielded search"? ____ yes ____ no
- c. What would you type into the **Enter your search term(s)** textbox? _____

Now please perform the search you planned in Step 18. TRY ONLY THIS ONE SEARCH.

Receive GILS Search Results.

- 19. How many "total hits" did GILS return? _____
- 20. Did you expect a "hit" that would obviously point you to the toll-free number? ____ yes ____ no
- 21. Do any "hits" appear to be relevant? ____ yes ____ no

246

247

If you answered "no" in either Step 20 or Step 21, why not? (Circle all reasonable possibilities.)

- a. I doubt if such a service exists
- b. I don't know enough about social security to come up with good search terms or to choose database(s)
- c. I'm unsure of "fielded searching" in GILS
- d. I don't think GILS would include telephonic information resources
- e. The phone number is probably on a website and therefore not duplicated in GILS due to the maintenance burden
- f. The Social Security Administration does not participate in GILS
- g. The toll-free phone service is probably too new to be included in GILS
- h. Other (please describe): _____

Please click on Netscape's **Bookmarks** button and select the first item: *Government Information Locator Service.*

Receive *Government Information Locator Service (GILS) Online via GPO Access.*

Click on the hypertext phrase **fielded search**; receive *Search GILS Records by Fields Online via GPO Access.*

In the scrollbox labelled **Make your selection(s)**, click on **All Records on GPO Access GILS Site.**

In the section labelled **Select one or more of the following fields (mandatory GILS core elements) to search:** click on the blue hypertext phrase **Control Identifier** (not the checkbox).

Please receive and read the definition of **Control Identifier**.

23. Is the definition, and how it fits into GILS, clear to you? _____ yes _____ no

249

248

Please click on Netscape's Back to return to Search GILS Records by Fields Online via GPO Access.

Click in the checkbox next to Control Identifier. Click Submit.

Receive GPO Access Government Information Locator Service--Search GILS Records by Relevant Term.

In the textbox labelled Control Identifier, type in USCCR0212 (where 0 is a zero). Please double-check your typing.

Click Submit; then take a moment to scan through GILS Search Results.

24. Keeping in mind that you have searched the "control identifier" field, whose contents "distinguish this locator record from all other GILS Core locator records", what is your reaction to the list of hits? (Circle one.)

- a. The record creators made an error
- b. I do not understand "control identifiers"
- c. The system has made an error
- d. One of these is something other than a GILS Core locator record
- e. The records are the same--one in English and one in Spanish
- f. I do not notice anything unusual about this search result
- g. Other (please specify): _____

Please click on Netscape's Bookmarks button and select the second item: Search the GILS Database.

Receive U.S. Environmental Protection Agency--Search the GILS Database.

Quickly scan the page.

251

250

25. Please check "yes" or "no" for the following:

	Yes	No
a. Are you surprised to find that EPA's Government Information Locator Service looks different from GPO Access' Government Information Locator Service?		
b. Do you expect the EPA's Government Information Locator Service to operate exactly like GPO Access' Government Information Locator Service?		
c. Do you think that EPA's Government Information Locator Service is part of GPO Access' Government Information Locator Service?		

26. What might you expect EPA's Government Information Locator Service to provide? (Please circle all that apply.)

- a. Statistics about the ozone layer
- b. A catalog of EPA publications
- c. Maps
- d. The full text of EPA regulations
- e. Descriptions of technical reports
- f. Congressional testimony on nuclear accidents
- g. Hotline phone number
- h. An order form for a radon-testing kit
- i. An abstract for a CD-ROM about nature
- j. Clinton's 1996 inaugural address
- k. Census data
- l. A list of Superfund cleanup sites
- m. A phone directory of EPA staff
- n. Images of the spotted owl
- o. Hotlinks to environmental activism websites

27. Please rate how confident you would feel using the following options presented on this screen:

	Certain	Unsure
Complete text		
Acronym		
Agency Program		
Local Subject Term		

In the textbox labelled **Complete Text**: please type the following search term:

wetlands AND "new england"

Click Submit Query.

Receive GILS Search Results.

Click on blue hypertext HTML for hit 1. Title: Index to the Wetland Educational Resources distributed by the New England regional office, EP.

Receive and SCAN the GILS record. DO NOT CLICK ON ANY HYPERTEXT IN THE RECORD.

28. Which of the following does this record describe?

- a. An index
- b. An educational "kit"
- c. Miscellaneous training items available separately
- d. Don't know
- e. Other (please specify): _____

29. How up-to-date are the described material(s)? _____

255

254

30. Would you assume the following to be “true” or “false”:

	True	False
a.	Information in EPA GILS is authored by EPA	
b.	Information in EPA GILS is the most current available	
c.	EPA is mandated by law to provide the information in GILS	
d.	EPA GILS describes every EPA publication	
e.	A duplicate of EPA GILS exists on GPO Access' GILS	
f.	I can find EPA GILS records by use of a web search engine such as <i>Yahoo!</i> , <i>AltaVista</i> , or <i>Lycos</i>	
g.	This GILS has only information resources of EPA Headquarters in Washington, and does not include regional offices	

31. How many GILS do you think may exist on the Internet today? (Circle one.)

- a. One for each Federal agency
- b. One for each branch of the government
- c. One on each website that has ".gov" as part of the address
- d. One for each broad subject area
- e. One for each type of information resource
- f. Only one
- g. I have no basis for guessing

Please take the remaining time to explore GILS on your own by using Netscape's Bookmarks 3 through 8 and then complete this final step:

32. Please rate your agreement with the following statements using a checkmark.

	STRONGLY AGREE	AGREE	DISAGREE	STRONGLY DISAGREE	NO OPINION
a. All agencies' GILS should be searchable together, from one website.					
b. There are not enough fields to search in a GILS database.					
c. All government documents on the Internet should be hotlinked from one electronic card catalog.					
d. It would be easier to search GILS records if they were grouped hierarchically by subject.					
e. I am unaware of the function or purpose of many Federal agencies.					
f. All GILS records should look alike.					
g. I will use GILS to locate government information in the future.					
h. GILS is an efficient service.					
i. GILS probably helps agencies manage their information resources.					
j. It is clear to me how agencies choose what to include in GILS.					
k. Public electronic access to government information is important.					
l. GILS is an improvement over microfiche and paper indexes.					
m. All GILS records should contain information in all fields.					
n. The quality of the records I examined varied widely.					

THANK YOU! Please hand this completed packet to an investigator, and join us for the debriefing. (We'll turn off your computer.)

259

253

**APPENDIX E-1
Summary of GILS Conference Survey Results**

APPENDIX E-1

Summary of GILS Conference Survey Results

Table of Contents

1.0. INTRODUCTION.....	1
2.0. DATA MANAGEMENT AND ANALYSIS	1
3.0. RESULTS	1
3.1. Characteristics of Respondents	1
3.2. Respondents Definitions of GILS	5
3.3. Assessment of Key Issues	7
3.4. Additional Comments Provided by Respondents.....	11
3.4.1. "Questions and Issues" Responses.....	12
3.4.2. "Complaints and Concerns" Responses	13
3.4.3. "Recommendation" Responses	14
3.4.4. "General" Responses	15
3.4.5. "Here To Learn" Responses	15
Table E1-1: Settings in Which Respondents Work.....	2
Table E1-2: Field in Which Respondents Work	3
Table E1-3: Involvement with GILS.....	3
Table E1-4: Respondent's Agencies' GILS Implementation	4
Table E1-5: Familiarity with GILS Documents/Policies.....	4
Table E1-6: Familiarity with GILS Documents/Policies (Reduced)	5
Table E1-7: Definition of GILS – Functions.....	6
Table E1-8: Type of Information in GILS	6
Table E1-9: Users of GILS	6
Table E1-10: GILS Coverage	7
Table E1-11: Key Issues	8
Table E1-12: Key Issues (Reduced).....	9
Table E1-13: Use of GILS	9
Table E1-14: Favorite Online Alternative to GILS	10

APPENDIX E-1

Summary of GILS Conference Survey Results

1.0. INTRODUCTION

The following provides an analysis and interpretation of the results of a survey questionnaire distributed to participants at the Second Annual GILS Conference (November 13-14, 1996). Approximately 300 people were scheduled to attend the conference. A total of 181 completed questionnaires.

The questionnaire was intended primarily to solicit from people who could reasonably be expected to know about GILS their assessment of a set of key GILS issues. In addition, the survey offered the opportunity to collect information to gauge respondents knowledge and awareness of GILS-related policy and technologies. The questionnaire was passed out early on the first day of the Conference and was intended to gauge respondents' assessments based on their existing knowledge rather than knowledge gained from the Conference.

2.0. DATA MANAGEMENT AND ANALYSIS

The questionnaire (see Appendix D-3 copy of survey instrument) included open- and closed-ended questions. Respondents were asked for minimal demographic information that would provide context as to the characteristics of the respondents and their responses. All data were entered into Access database and statistics generated via Excel. Once the data were entered, 20 randomly selected questionnaires (approximately 10%) and their corresponding database entries were reviewed for accuracy.

Simple descriptive statistics were produced for the data from the closed-ended questions. For the open-ended questions, content analysis was performed to categorize and understand responses. Preliminary coding categories were developed based on the actual data rather overlaying a coding scheme on the data.

3.0. RESULTS

The following sections report on the results of the survey. Section 3.1 Characteristics of Respondents addresses the demographic information provided.

3.1. Characteristics of Respondents

Questions 2 through 7 asked respondents for information about where they work, what type of work they do, their involvement with GILS, and their familiarity with GILS policies and documents.

Settings in Which Respondents Work

Question 2 asked respondents were asked to identify the *setting* that best described where they worked. Four named choices were offered with a fifth choice of "Other." Table E1-1 summarizes the data from Question 2. Given the venue and focus of the GILS Conference, it is not surprising that the vast majority of respondents work in Federal agencies.

Table E1-1
Settings in Which Respondents Work

Setting	N	%
Federal Agency	145	78%
State or Local Government	17	9%
Private Non-Profit Organization	6	3%
Private For-Profit Organization	6	3%
Library	4	2%
Contractor to agency	2	1%
Other	5	3%
TOTAL	185*	99%**

* Respondents marked multiple choices so N exceeds the 181 completed surveys.

** Total does not equal 100% due to rounding.

The category of "Other" included the following respondent-generated answers to identify further their settings:

- Canadian Government
- Commonwealth (Australia)
- Consultant to all of the above
- Employee Commission
- European Research Agency.

Years of Experience

Question 3 asked respondents for the years of experience they had in this setting. The average number of years was 15 years with a maximum of 37 years reported. The mode was 20 years, which indicates that the majority of respondents are not novices in their work settings. In fact, if one extrapolates an "age" of respondent based on their years in the setting, one can assume that the majority of respondents were between 40-50 years old (assume work life beginning at 21 plus 20 years in the setting).

Field in Which Respondents Work

Question 4 asked respondents to identify the *field* in which they worked. As in question 2, respondents were offered twelve named choices, with an additional choice of "Other." Table E1-2 summarizes the data from Question 4.

Table E1-2
Field in Which Respondents Work

Setting	N	%
Records Management	67	34%
Information Resources Management	35	18%
Library/Information Center	29	15%
Public Information	19	10%
Program Office/Project Management	10	5%
Computer Systems	8	4%
Archives	7	4%
Chief Information Office	6	3%
Legal/Legislative	4	2%
Software Developer	3	2%
Sales/Marketing	3	2%
Research and Development	1	1%
Procurement/Contracting	0	0%
Other	8	4%
TOTAL	200*	104%**

* Respondents marked multiple choices so N exceeds the 181 completed surveys.

** Total does not equal 100% due to rounding.

The category of "Other" included the following respondent-generated descriptions to identify further their field of work:

- Administrative services, including document management, mail services, etc.
- Advocacy
- Management Analyst
- State-wide Planning
- Trainer, Outreach Analyst
- Web Entrepreneur
- Webmaster for Office (new position).

Respondent Involvement with GILS

Question 5 asked respondents about *their* involvement with GILS. Six named choices were available to the respondents, with an additional choice of "Other. Table E1-3 summarizes respondents answers. The majority of respondents have some involvement with GILS.

Table E1-3
Involvement with GILS

Involvement	N	%
Implementor	81	39%
GILS Record Creator/Maintainer	46	22%
User	39	19%
Policymaker	19	9%
Technical Standards Developer	5	2%
Information Reseller	4	2%
Integration/Interoperability	2	1%
Trainer/User Support	2	1%
Other	8	4%
TOTAL	206*	99%**

*Respondents marked multiple choices so N exceeds the 181 completed surveys.

**Total does not equal 100% due to rounding.

The category of "Other" included the following as stated by the respondents to identify further their involvement with GILS:

- Curious about GILS, possible future user
- I wear more than one categorical hat
- Information management consulting
- Information systems developer
- Rabble rouser
- Software tools
- Technical support
- Trying to find out what GILS is and its requirements.

In terms of the audience for which Conference was intended, the majority of respondents were involved in one way or another with GILS. The specifics of that GILS involvement was not addressed by the survey. For example, involvement as "user" did not specify the type of use nor the extent of GILS use by the respondents.

Agency Involvement with GILS

Question 6 asked respondents to describe their agencies' involvement with GILS and their implementations. Table E1-4 summarizes the results from this question.

Table E1-4
Respondent's Agencies' GILS Implementation

Agency Implementation	YES		NO		D/K		TOTAL	
	N	%	N	%	N	%	N	%*
Operates a GILS Server	66	46%	60	42%	16	11%	142	99%
Submits Records to Others	59	42%	59	42%	24	17%	142	101%
Part of Larger GILS Effort	69	51%	47	35%	20	15%	136	101%

* Total does not equal 100% due to rounding.

Respondent Familiarity with GILS Documents & Policies

Question 7 attempted to gauge respondents current knowledge (prior to the Conference) of relevant GILS policies and documents. Respondents identified their familiarity on a Likert Scale with 1 = Very Familiar and 5 = Not Familiar. Table E1-5 summarizes the results from this question.

Table E1-5
Familiarity with GILS Documents/Policies

GILS Documents/Policies	1		2		3		4		5		Blank		Total	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%*
Federal GILS Policies	45	25	41	23	40	22	22	12	30	17	3	2	181	101
Agency's GILS Policies	49	27	32	18	24	13	19	10	28	15	29**	16	181	99
NARA's Guidelines	40	22	42	23	32	18	25	14	38	21	4	2	181	100
OMB Bulletin 95-01	40	22	46	25	30	17	28	15	34	19	3	2	181	100
Z39.50 Standard	7	4	31	17	45	25	35	19	57	31	6	3	181	99
FIPS No. 192	17	9	24	13	30	17	45	25	61	34	4	2	181	100
PRA 1995, GILS Section	45	25	36	20	45	25	28	15	25	14	2	1	181	100
The World Wide Web	72	40	49	27	30	17	17	9	11	6	2	1	181	100

* Total does not equal 100% due to rounding.

** Due to high number of conference attendees not employed by a federal agency.

Another perspective on the answers to Question 7 is to collapse the cells for the Likert Scale choices into one for Familiar (comprising 1 and 2) and one for Not Familiar (comprising 4 and 5). This isolates the cells for Likert Scale choice 3. Table E1-6 presents this information.

Table E1-6
Familiarity with GILS Documents/Policies (Reduced)

GILS Documents/Policies	Familiar		3		Not Familiar		Blank		Total	
	N	%	N	%	N	%	N	%	N	%*
Federal GILS Policies	86	48	40	22	52	29	3	2	181	101
Agency's GILS Policies	81	45	24	13	47	26	29	15	181	99
NARA's Guidelines/Record creation	82	45	32	18	63	35	4	2	181	100
OMB Bulletin 95-01	86	48	30	17	62	34	3	2	181	101
Z39.50 Standard	38	21	45	25	92	51	6	3	181	100
FIPS No. 192	41	22	30	17	106	59	4	2	181	100
PRA 1995, GILS Section	81	45	45	25	53	29	2	1	181	100
The World Wide Web	121	67	30	17	28	15	2	1	181	100

* Total does not equal 100% due to rounding.

When one looks at Table E1-6, it becomes evident that the majority of respondents do not have familiarity with some of the basic documents and policies that underpin the GILS effort. Less than half of the respondents stated familiarity with OMB 95-01. The responses to this question raise the issue of the extent and nature of training and education on GILS that has been available. Since Z39.50 is a key aspect of GILS, it is worrisome that only approximately 20% of the respondents were familiar with either the standard or the Federal Processing Standard, FIPS No. 192, that mandates the use of Z39.50 for agency GILS implementations.

3.2. Respondents Definitions of GILS

The researchers were interested in determining what respondents thought GILS was. Question 1 on the survey asked respondents to provide their definition of GILS. This was an open-ended question and respondents provided a wide range of answers to the questions. Content analysis was performed on the data and the answers were grouped into four categories based on the primary features or aspects of GILS contained in the definitions:

- Describing GILS from the perspective of its *functions*
- Describing GILS from the perspective of the *types of information* comprising GILS
- Describing GILS from the perspective of various *potential users* of GILS
- Describing GILS from the perspective of the *coverage* of GILS.

Respondent definitions often addressed more than one feature or aspect.

Within each of these categories, the content analysis revealed additional details that were also coded and grouped. The following tables (Table E1-7 through E1-10) summarize the results of the content analysis on the data provided in Question 1. For each table, the percentage of respondents describing that perspective on GILS is given.

Table E1-7
Definition of GILS -- Functions

84% of respondents' definitions characterized GILS in terms of its function(s):

Type of Function	N	%
Finding Aid	97	64%
Access	47	31%
IRM	4	3%
Collect	3	2%
Control	1	1%
TOTAL	152	99%*

* Total does not equal 100% due to rounding.

The specific "functions" categories were defined based on the occurrence of words in the respondents' definitions:

- Finding Aid -- defined by words such as "card catalog," "index," "pointers," etc.
- Access -- defined by words such as "provide access to," "retrieve information," etc.
- IRM -- defined by words such as "managing resources," "records management," etc.
- Collect -- defined by words such as "agencies 'collect' information via GILS"
- Control -- defined by words such as "agencies 'control' information via GILS."

Table E1-8
Type of Information in GILS

38% of respondents' definitions characterized GILS in terms of the type(s) of information it comprises:

Type of Information	N	%
Publications	18	26%
Resources	18	26%
Systems	12	17%
Records	16	23%
Services	5	7%
TOTAL	69	99%*

* Total does not equal 100% due to rounding.

Table E1-9
Users of GILS

24% of respondent definitions characterized GILS in terms of the type of user(s):

Type of Users	N	%
Public	32	74%
Agency	4	9%
Private	2	5%
Government	1	2%
Library	1	2%
Managers	1	2%
Researchers	1	2%
World	1	2%
TOTAL	43	98%*

* Total does not equal 100% due to rounding.

Table E1-10
GILS Coverage

13% of respondent definitions characterized GILS in terms of its coverage:

Coverage	N	%
Federal government information (no qualifier mentioned)	7	29%
Important/major/prime information	4	17%
Executive information	1	4%
Electronic information	1	4%
Useful information	1	4%
Other	10	42%
TOTAL	24	100%

The category of "Other" in Table E1-10 includes the following limitations to GILS coverage:

- A basic replacement and improvement to requesting information from Pueblo, CO -- you can find all agencies with information on topic
- [primary] systems of records
- Certain federal holdings
- Information federal agencies choose to make available
- Government services policy procedures information
- Public records to patrons of the service
- Records federal agencies are creating
- Technical knowledge gained through research
- All of IRS systems
- Information for govt. agencies to complete daily duties

Additionally, seven (7) respondents indicated some aspect of GILS related to its centralization (2) or decentralization (5). Five (5) respondents mentioned standards including two (2) respondents specifically mentioned Z39.50.

Other responses were difficult to categorize, but contain interesting and sometime quite honest comments about GILS:

- A concise instrument to facilitate the use of government information
- Networked magic black box for finding universal government information
- Providing information to private agency, schools, etc., on data related to that agency, company, etc.
- Roads signs identifying communities of interest on the information highway
- This is my first GILS conference. I was not given any background on GILS
- Wish I knew.

Overall, the answers to this question reveal a wide diversity in how people currently understand and are able to define GILS. Two key aspects of GILS does come through in the answers -- that its "users" are considered to be "public users," and its primary characteristic is as a "finding aid."

3.3. Assessment of Key Issues

As noted in the introduction, the primary purpose of this survey was to get input for the study on issues related to GILS. The researchers identified through the literature review, early data collection activities, and discussions with people knowledgeable about GILS a set of statements that reflected a preliminary list of issues related to GILS. Respondents were asked to state their agreement/disagreement with thirteen (13) statements and true or false for four (4) statements. Table E1-11 summarizes the data from the thirteen 13 statements. For those statements, respondents

could choose on a five-point Likert Scale (1 = Strongly Agree and 5 = Strongly Disagree) or make the choice "Don't Know" (D/K). Table E1-12 collapses cells for 1 and 2 together and for cells 4 and 5 together to indicate more general senses of agreement/disagreement with the statements. Table E1-13 summarizes the data from the true/false statements.

**Table E1-11
Key Issues**

Key Issue	1		2		3		4		5		D/K		Blank		Total	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
A purpose of GILS is to improve public access to government information	128	71	33	18	14	8	0	0	2	1	4	2	0	0	181	100
Every agency Web homepage should have a link to the agency's GILS	97	54	46	25	15	8	2	1	3	2	15	8	3	2	181	100
GILS should be expanded through out the Federal government	73	40	43	24	24	13	5	3	5	3	25	14	6	3	181	100
A purpose of GILS is to help agency officials better manage agency information	51	28	49	27	32	18	19	10	10	6	16	9	4	2	181	100
More work needs to be done on the technical standards to support GILS operations	43	24	39	22	31	17	10	6	2	1	50	27	6	3	181	100
I am able to describe GILS accurately and fully to others	37	20	44	24	53	29	19	10	20	11	3	2	5	3	181	99
GILS has improved public access to Federal information	32	18	31	17	40	22	18	10	3	2	52	29	5	3	181	101
GILS is so valuable that, if it didn't exist, we would have to create it	23	13	34	19	50	28	30	17	11	6	27	15	6	3	181	101
There is adequate policy guidance to direct the development and operation of GILS	12	7	27	15	46	25	34	19	12	7	43	24	7	4	181	101
The appropriate technology is readily available for implementing GILS	11	6	32	18	48	27	20	11	13	7	50	27	7	4	181	100
There is adequate technical implementation guidance for development and operating GILS	9	5	27	15	47	26	28	15	17	9	49	27	4	2	181	99
The World Wide Web reduces the need for GILS	9	5	20	11	28	15	49	27	39	22	29	16	7	4	181	100
GILS records represent the complete information resources of an agency	7	4	5	3	20	11	53	29	69	38	22	12	5	3	181	100

* Total does not equal 100% due to rounding.

Table E1-12
Key Issues (Reduced)

Key Issue	Agree		3		Disagree		Blank		D/K		Total	
	N	%	N	%	N	%	N	%	N	%	N	%*
A purpose of GILS is to improve public access to government information	161	89	14	8	2	1	4	2	0	0	181	100
Every agency Web homepage should have a link to the agency's GILS	143	79	15	8	5	4	15	8	3	2	181	101
GILS should be expanded through out the Federal government	116	64	24	13	10	6	25	14	6	3	181	100
A purpose of GILS is to help agency officials better manage agency information	100	55	32	18	29	16	16	9	4	2	181	100
More work needs to be done on the technical standards to support GILS operations	82	45	31	17	12	7	50	28	6	3	181	100
I am able to describe GILS accurately and fully to others	81	45	53	29	39	22	3	2	5	3	181	101
GILS has improved public access to Federal information	63	35	40	22	21	12	52	29	5	3	181	101
GILS is so valuable that, if it didn't exist, we would have to create it	57	31	50	28	28	23	27	15	6	3	181	100
The appropriate technology is readily available for implementing GILS	43	24	48	27	33	18	50	27	7	4	181	100
There is adequate policy guidance to direct the development and operation of GILS	39	22	46	25	46	25	43	24	7	4	181	100
There is adequate technical implementation guidance for development and operating GILS	36	20	47	26	45	25	49	27	4	2	181	100
The World Wide Web reduces the need for GILS	29	16	28	15	88	49	29	16	7	4	181	100
GILS records represent the complete information resources of an agency	12	7	20	11	122	67	22	12	5	3	181	100

* Total does not equal 100% due to rounding.

Questions 22 - 25 asked respondents about their use of GILS. Respondents marked True or False for their answers to these questions. Table E1-13 summarizes responses to these questions.

Table E1-13
Use of GILS

Use of GILS	True		False		Total	
	N	%	N	%	N	%
I find useful information when I use GILS	84	46%	97	54%	181	100%
I often find links to GILS on the Web	53	29%	128	71%	181	100%
I often refer people to GILS when providing information	48	27%	133	73%	181	100%
I search GILS several times per day in my everyday work	6	3%	175	97%	181	100%

A large majority of respondents neither use GILS frequently nor do they refer people to GILS for finding information. A smaller majority (54%) do not find useful information when using GILS. Responses to these questions raise the question as to the utility people currently see GILS as providing.

Alternatives to GILS

Along with the questions related to what users think about GILS as reflected in answers to questions 22-25, another question (Question 8) asked users to identify their favorite online alternative (either Web-based or other) to GILS they use when trying to locate government information. This was an open-ended questions, and respondents answered ranged from "The Web" to specific "agency homepage(s)" or agency locators such as "GPO Access." While many respondents noted more than one alternative, only the first listed was counted in Table E1-14, which shows a summary of responses to Question 8. A second part to Question 8 asked respondents to provide reasons why they liked the source listed in Question 8. The intention was to identify salient features or aspects of those sources to provide a primitive sense of "user requirements" for changes to GILS. In general, and with some notable exceptions, the responses to the second part of the question were not useful for the intended purposes. Table E1-14 is a list of some of the more interesting features of these alternative sources.

Table E1-14
Favorite Online Alternative to GILS

Alternative Online Resource	N	%
Web	32	18%
Alta Vista	14	8%
Agency Homepage(s) (includes DefenseLink)	13	7%
GPO Access	11	6%
None**	7	4%
Fedworld	6	3%
Yahoo!	6	3%
No Access**	5	3%
White House website	4	2%
(unspecified) Search engine	4	2%
GPO Monthly Catalog	3	2%
Library of Congress (includes Thomas)	3	2%
Villanova	3	2%
Lexis/Nexis	2	1%
Lycos	2	1%
Netscape	2	1%
Webcrawler	2	1%
EPIC	1	1%
Excite	1	1%
Gopher	1	1%
Library	1	1%
metacrawler.cs.washington.edu	1	1%
Mosaic	1	1%
Nonprofit Advisor	1	1%
NSDI/FGDC	1	1%
Telnet	1	1%
Usenet	1	1%
WWW.strategics.ca	1	1%
Blank	51	28%
Totals	181	105%*

*Total does not equal 100 due to rounding.

**"None" includes responses such as "don't use" and "not familiar with any"; "no access" includes responses such as "not available" and "we're not online".

Respondents noted the following reasons for favoring the GILS alternatives listed in Table E1-14:

- Ease of use
- Access to multiple/a variety of official databases
- Currentness
- Pointers to nearest library
- Helpful user's guide
- Extensive index
- Absence of advertising
- Speed
- Amount of information provided
- Clear organization
- Convenience/familiarity/lack of knowing alternatives
- Keyword search capability
- Comprehensiveness
- Breadth/depth of subject matter
- Support
- Links to other government agencies/resources
- Seeing a "better perspective (context) of" information
- Size
- Relevance of information retrieved
- Synergy
- Personal involvement in development/maintenance
- Target audience
- Standardization
- Ability to search by document title
- Variety of approaches to information organization
- Wide availability
- Graphics development and presentation
- Efficiency
- Intuitiveness
- Support of libraries, public information , and preservation
- Predictability
- Results are clear
- Simplicity
- Familiarity
- Support by federal agencies
- Degree of organization
- Definitiveness
- Capable of complex searching
- Interface
- Solid data
- Script-ability to allow access for other programs.

An important design question for GILS is whether GILS can be modified to address and accomplish these criteria and expectations of users.

3.4. Additional Comments Provided by Respondents

The survey included a final question (Question 26) that asked respondents to offer any additional comments, pose questions, identify key issues, or discuss topics not mentioned in the questionnaire. The intention of this question was to probe for respondents' ideas that the GILS Team would consider in subsequent data collection activities.

This was an open-ended question, and respondents provided a wide range of responses. A preliminary categorization placed the questions into several broad groups:

- Questions and Issues
- Complaints and Concerns
- Recommendations
- General
- "Here to Learn"

3.4.1. "Questions and Issues" Responses

- Where will the ""front door"" be for the federal GILS?
- How important is z39.50 to implementation of GILS? Doesn't web circumvent need for that protocol?
- I find a void in clearly articulated description of what GILS is, what my responsibilities are, and how we are to implement GILS.
- Questions: in this age of diminished resources, my agency cannot afford automation systems or personnel to perform mission related [to] automation support. In this environment of exceptionally constrained resources (\$) how can the agency afford another ""unfunded mandate""? How does the agency choose which ""unfunded mandate"" to comply with?" Is this truly the wisest use of our dollars? Isn't it GILS (etc.) Versus mission accomplishment!
- DOD GILS: we are directed by OSD, PAIGE to use DTIC as our vehicle to put our records out to the GILS which they carry out very well but we (including DTIC) have been directed to use the DIST [D.I.S.T.]. Nice easy step for DTIC but inaccurate data is put out to GILS because of this nice and easy step. #19 [web reduces need for GILS] I answered strongly agree but another [question could be] if not directed to have put out the GILS info would your agency have done this on their own?
- Relationship between GILS and NTIS FEDRIP (federal database of R&D research in progress). Relationship between GILS and critical technologies institute (CTI) funded by OSTP database of government R&D. Radius (R&D in the U.S. government). How differ?
- Relationship of state-level GILS with federal GILS?
- I was assigned to be the creator of GILS records for our agency. However, I still do not fully understand how it should be created. For instance, if an agency has much of or the majority of the its information that is requested by the public on the Internet: is it feasible to conclude that the homepage of the WWW would meet the GILS requirement?
- Issue: what are personal incentives for maintaining & for using GILS?
- Should there be a single web site for federal government information? Why do GILS records have to comply with z39.50 and FIPS 192? Does the public care?
- How can we capitalize on z39.50 base and link U.S. fed. GILS records to huge base of library catalogs? How can we make sense of the z39.50 thing?
- The application of EDMS technology to GILS core elements- especially as vendors tackle the use of z39.50 with their products - meta data, search engines, profiles, etc.
- 1. What is the degree of adherence to GILS within the fed. Govt.? 2. What is the commitment within depts & agencies? 3. Is z39.50 being used at all or are GILS sites primarily web-based? 4. What is being done to promote GILS when depts and to the public? 5. Do depts. Use guidelines to determine which inform. Resources to describe in a GILS record? 6. What are the obstacles to GILS implementation? 7. What is the status of GILS in other natl gov'ts? 8. How can Netscape catalog server be used to support GILS?
- We have heard about GILS becoming global. I'm more interested in how GILS will be adapted at state and local level, even by non-governmental entities. To what degree is GILS not the WWW? (re question 19)
- Key issue: identifying potential customers, not just current users. Who uses GILS besides agencies & federal depository librarians?
- Issue: keeping the GILS info current, accurate, up-to-date
- How to begin to establish a state GILS. How does one decide which agencies to approach to describe their info resources & put them online?
- Discussions of implementing GILS in our state government yielded a variety of responses, to wit: yes, it's necessary, but is the federal model one to emulate? Libraries asked whether GILS records could be MARC records; there were questions of how to collect GILS info (centralized or delegated), and how best to capture subject access to government info holdings (LCSH or just let the agencies do it?). These are not roadblocks, but represent some of the obstacles GILS needs to overcome.
- Getting greater commitment within agencies for GILS - resources needed - support by agency heads as a priority activity.
- Key issue: procedures for maintaining improvements to online record development & maintenance.
- Effective date of FIPS PUB 192? NARA guidelines on mandatory fields - is this info available?

3.4.2. "Complaints and Concerns" Responses

- The quality of GILS varies widely from agencies to agencies.
- Hope some of the GILS will be linked to the actual document in some way -- hope less technical descriptions. -- some are just doc #'s which do not make sense to lay persons.
- Access to one of a kind archival information.
- Volume provides cost effective indexing [for archival information].
- Links to document management systems.
- The naming, meaning, and organization of most of the GILS records fields are much too obtuse for the casual user (and not record submitters). Currently, we have found this to be our biggest handicap to acquiring new records.
- I have never been able to locate a resource to advise on how to implement GILS. Our agency submitted a requisition to GPO to put up our system of records. GPO has never billed us and when we inquired, we aren't able to locate anyone in GPO who knew anything about GILS. Our customer service rep at GPO could not help either.
- GILS is being shoe-horned into/used for purposes that are totally inappropriate - records disposition. GILS is either about important govt information, or it isn't.
- 1) I am concerned that there are several initiatives (such as GILS & FGDC) that are duplicating efforts (or at least having that functional result) in storing govt info. 2) I am concerned that storing data in GILS is a separate activity, hence adding to workload. Data (including meta data) needs to be gathered and stored at a single point, as a result of the initial data collection/storage activity(s). In other words, the capture and storage needs to be automated as an offshoot of the primary activity. 3) the level of the records I've seen is so high it's almost meaningless.
- There is no way for a member of the public to search across all GILS records. How can a member of the public find an agency's GILS records?
- 1. While the process of coding GILS elements is structured, the penetration of GILS into the whole population of information is not controlled. 2. Dependence upon keyword improves precision, but not recall. As more general employees assist with coding GILS entries, the variety of cataloging terms destabilizes accuracy.
- 1) agency politics among different program areas has been roadblock in implementing GILS.
- Requirements to put expensive publications on GILS concerns many employees in our office. We do not have the staffing to hire people to handle new requests, but we also don't have the equipment to put all this info (publications) online.
- Too much effort into policy, not enough into the technology
- A problem for some agencies is that the necessary GILS people have no web links because of budget constraints.
- The value of GILS is minimal because the descriptions do not always accurately reflect the information available. In most cases, you cannot determine if the publication/information referenced is indeed what I want. The records management capabilities are nonexistent. This aspect of GILS needs to be re-thought. The national archives is not ready to accept GILS records in lieu of a schedule. In addition, the GILS record does not provide the information needed by NARA to appraise records--therefore generating more paperwork.
- I think the WWW is a far superior way to get federal info. GILS is just more bureaucratic red tape.
- Conflict between privacy act notice requirements and duplicate/different GILS requirements still avoided & unresolved. Ditto conflict between NARA scheduling & RM paperwork & process requirements v. GILS requirements - differing levels of GILS entries inconsistent & confusing. Scarce agency resources now devoted to trying to maintain & reconcile duplicate information management databases. Implications of EFOIA uncertain despite Katzen's rosy glow.
- The strong tendency in federal agencies to move slowly toward web access for their info (whether or not they provide links to their GILS). The extreme level of aggregation of records. The lack of ability to search easily/transparently across agencies. The lack of a search thesaurus. The lack of access to what is pointed to (esp. As concerns numeric databases) for the most part. The lack of vision (forward or peripheral) from OMB, the admin.

- Examine how OMB placed the requirement, and how that affected agencies implementations. If GILS is so important, where are the resources (\$, information, tools) that should have been made available? Another strategy other than an OMB bulletin might have had a different effect/outcome.
- My agency's IRM staff locks out the records creator and either takes over the final stage of the project, or deems the information project non-helpful, and torpedoes years of information gathering
- Most users who I come in contact with are very intimidated by GILS because it is very foreign to them -- at times ""too formal." They are not aware of the value of this service. I also think there needs to be a [single] place to go to get GILS records. Not having one place to find GILS records is very rough on users.
- When will OMB recognize & provide reasonable resources for GILS implementation, maintenance, and improvement?
- GILS--not user friendly. Hard to search. Explain how to search - i.e., privacy act systems of records.
- Perceived redundancy w/ NSDI/FGDC is a major stumbling block to GILS acceptance
- Costs & impact on small agencies without staff or resources to meet unfunded mandates. Guidance on protection of classified, client privileged information as relates to GILS and addition of records schedules/disposition. Difficult to get to specific sources, often get irrelevant data, search sources don't lend themselves to precise retrieval of information.
- Speaking as a citizen, client, customer, and subject (see Mintzberg's article on ""governing"" HBR, 1996, management...) I regret saying that the "bloom is off the rose." GILS sucks. It should meet the real world. Z39.50 is snake oil.

3.4.3. "Recommendation" Responses

- Library established standards should be used for GILS.
- Would like more information on use of GILS for records management.
- Greater awareness of GILS needed.
- Compare/contrast GILS vs. GPO MOCAT. Legislation should be enacted to develop GILS for legislative and judicial branches of government. Improved guidance to agencies is necessary; tying GILS to agency mission/dissemination.
- Intergovernmental cooperation - federal, state, local participation [and] coordination in disseminating public records/information. OMB as a record management tool.
- Can we work on integrating historical information and older publications with pointers to source (i.e., dept. Libraries, NARA, NTIS)? Can we work on expanding GPO Access or other source to a government web page with pointers to each agency - include judicial with congressional and executive branches? What about education packets for DIST. To school labs on using GILS.
- Never thought I'd say this but can't someone write a GILS how-to-do-it manual in plain English to take away much of the mystery?
- How do you explain GILS to upper level politicians & policy makers?
- It would seem that a natural step for GILS would be a migration to a www/x.500 platform for GILS.
- Need to ensure that multicultural perspective is built into the design of GILS - also diversity in terms of educational level, age, groups, etc.
- OMB meta data should be used in developing an electronic records keeping system for consistency throughout govt.
- OMB doesn't seem to be a priority within my agency. I'm struggling with trying to educate my management in this area. My agency seems to be more enamored with technology rather than using the technology as a tool to disseminate information. I would like to see OMB & NARA being more proactive with educating agencies. I would also like to see OMB & NARA request implementation plans from agencies. Maybe this will jump start the effort.
- Need better direction to department and agency heads
- Need more universal compliance w/ standards, level of records. Need more easy access to records & provide html links to actual resources available electronically. More agencies need to use as record management tool. (complete).
- Normalize GILS data fields

- OMB input should be updated in a timely fashion. OMB input should be done once with no [conversion?] Needed. OMB needs to prevent privacy act violations.
- Develop hot links from records to data - MARC 856
- #16 [homepage link to GILS], vice versa is even more important; if the resource is out on the web page, there should be a hotlink from the corresponding GILS record.

3.4.4. "General" Responses

- It is positive that so many agencies are using the government printing office to fulfill their GILS obligations.
- Don't have access to Internet and other agencies GILS topics. Also, I really never comprehended purpose of GILS.
- I think that some of the info on GILS may be elsewhere on web.
- Lists of lists are basically useless. We need actual information, but providing this is costly- if agencies don't commit budget and people resources, GILS will be at best half-baked. Customers will always take the easiest route to information, which is to call up a human being who will track the info for them.
- GILS v. Web. There are valid arguments on both sides. The web seems to be infinitely more democratic. I wonder if GILS' time hasn't already come and gone.
- GILS will never be what it should be until all federal agencies are using EDMS software to capture documents and meta data through out their life cycles. When that occurs, and only then, will agencies be able to make information readily available to the public in whatever form they want it. Only then will the vision of GILS be realized.
- This questionnaire seems geared toward federal employees who use the federal GILS. I am not in that category, so I am not sure how useful my responses will be to you. #11 [adequate policy guidance]: whose GILS? Yours? Or my potential GILS?
- I believe the more important future of GILS relates to the management of information as opposed to the location or access to information. The web appears to be more functional for those purposes.
- FOIA interface. Legal/ethical issues. Information integrity/security. Information updates. Timeliness of information. Empowerment vs. Elimination of employees. Training initiatives. I am very interested in participating in a focus group [on] FOIA/policy analysis--assessing online GILS usage. Marcia Krug: w-(300) 413-0610, h-(300) 924-5104. Call me!
- As an information specialist in a government agency, I increasingly refer users to GILS.
- I think that homepage links may be useful, but often may be a big time/ bandwidth waster. If standards for relevance & completeness of access were developed and enforced, then GILS links to homepages could be useful. Content rules!
- Regarding #17 [improved public access], I'd say GILS will; I doubt it has, yet. Re: #22-25, I'm an old fogey who hasn't yet touched the web. Don't judge by me!

3.4.5. "Here To Learn" Responses

- I am not a GILS. Attendance at this conference will greatly educate me on the purpose and use of GILS.
- I would like to be involved in implementing GILS for my agency (Air Force)
- I am here to learn about GILS.
- Up until today I have not used GILS but plan to explore it when I return to my office. It sounds great!
- #22-#24 [search GILS several times per day, refer people, find useful info], not very familiar with GILS. My purpose here is to find out more info about it and how to access it.
- Its not very easy for me to honestly complete this survey because I really am not familiar w/ GILS, except that it is a Govt. Info. Locator service. I can only guess at some of these questions.
- A year and a half ago, I was asked to prepare a GILS record for my agency. It was duly sent up some chain of command, no doubt into integrated with others within our department (HHS) and I never heard another thing about it or GILS. I am a frequent user of the WWW and have some responsibilities for creating and maintaining our extensive website - which is hugely successful and much accessed. So what is the relationship of all this with GILS? I came to this conference primarily out of curiosity--because of my experience 1 1/2 years ago.

- I don't use GILS so don't count these answers. I have not been an active participant during the early work on my state GILS project; thus this conference is an introduction for me before I join the team during the prototype phase. Thus I think my answers should be discarded and not included in the results.
- I have not used GILS yet.
- I'm new at this, so I haven't used GILS yet.
- I have come to this conference representing the Australian government. In one of the reports prepared we recommended that Australia adopt a model based on GILS in order to provide access to govt info. I would like to keep in touch with you
- In process of finding out about GILS.
- I have not knowledge of GILS, whatsoever. I was told to attend this conference on Tuesday, in the place of my supervisor.

277

**APPENDIX E-2
Record Content Analysis
Findings, Discussion, and Recommendations**

APPENDIX E-2

Record Content Analysis

Findings, Discussion, and Recommendations

Table of Contents

1.0. Introduction	1
1.1. Organizational of Material	1
2.0. Method Overview	3
3.0. Data Summaries	3
3.1. Accuracy	3
3.2. Completeness	4
3.3. Resource Profile	4
3.4. Serviceability	5
3.5. Examples of High-Quality Records From Sample	6
4.0. Findings, Discussions, and Recommendations	7
4.1. Accuracy	7
4.2. Completeness	8
4.3. Resource Profile	23
4.4. Serviceability	29
5.0. Summary of Recommendations	33
5.1. Agency Level	33
5.2. Standards Level (Revision of the NARA <i>Guidelines</i>)	33
6.0. Further Research	34
7.0. Conclusion	35
Table E2-1 Organization of Results	2
Table E2-2 Aggregation Semantics	25
Attachments E2-1a through E2-1d: Examples of High-Quality Records From the Sample	37
Attachments E2-2a through E2-2d: Aggregation Code Examples	51

APPENDIX E-2

Record Content Analysis

Findings, Discussion, and Recommendations

1.0. INTRODUCTION

Content analysis of GILS records served three purposes: to assess records' quality in terms of completeness and accuracy; to explore the relationship of selected characteristics of records and serviceability in networked information discovery and retrieval (NIDR); and to develop recommendations for future application or adaptation of the method.

More than 3500 instances of metadata were evaluated for incidence and/or content, and entered into a database for subsequent coding and analysis. This appendix presents the results of that analysis, along with a discussion of implications and recommendations. In addition, the investigators maintained a log of areas for further research (see Section 8.0) that may be utilized by system developers, specification and procedures writers, and those with direct responsibility for GILS record quality.

1.1. Organization of Material

Section 3.0 Data Summaries aggregates significant results in terms of the analysis objectives. Detailed results of the analysis, Section 4.0 Findings, Discussion, and Recommendations, are presented in four major categories, which were assigned alphabetic codes. "Accuracy (A)" concerns the incidence of errors. "Completeness (C)" includes data concerning GILS record element utilization and values. "Resource Profile (P)" includes findings concerning general characteristics, such as aggregation and objects represented, for the sample population. The final section, "Serviceability (S)," presents findings relevant to record effectiveness in NIDR and user convenience. Further, each category code was coupled with a numeric code that reflects the order of data collection; under "Completeness" the data are additionally sorted in preferred element display order [per *Federal information processing standards publication 192, Application Profile for the Government Information Locator Service (GILS)* (National Institute for Standards and Technology, 1994) and *The government information locator service: Guidelines for the Preparation of GILS Core Entries* (National Archives and Record Administration, 1995a)] for ease of reference.

The following table offers a summary of the organizational scheme used in Section 4.0 Findings, Discussion, and Recommendations.

Table E2-1
Organization of Results

<u>CATEGORY AND CRITERIA</u>	<u>SECTION 4.0</u>
	<u>SUBSECTION</u>
	<u>(ITEM</u>
	<u>CODE)</u>
ACCURACY	4.1
FORMATTING ERRORS	A1
SPELLING AND TYPOGRAPHICAL ERRORS	A2
COMPLETENESS.....	4.2
NUMBER OF ELEMENTS PER RECORD.....	C1
PRACTICE OF PRESENTING "BLANK" ELEMENTS	C2
<u>UTILIZATION AND SELECTED CHARACTERISTICS OF "MANDATORY" ELEMENTS</u>	<u>C3</u>
TITLE.....	C3.1
ORIGINATOR	C3.2
LOCAL SUBJECT INDEX TERMS—"US FEDERAL GILS".....	C3.3
ABSTRACT	C3.4
PURPOSE.....	C3.5
AGENCY PROGRAM	C3.6
AVAILABILITY-DISTRIBUTOR	C3.7
AVAILABILITY-ORDER PROCESS	C3.8
SOURCES OF DATA	C3.9
ACCESS CONSTRAINTS.....	C3.10
USE CONSTRAINTS	C3.11
POINT OF CONTACT	C3.12
SCHEDULE NUMBER	C3.13
CONTROL IDENTIFIER	C3.14
RECORD SOURCE	C3.15
DATE OF LAST MODIFICATION	C3.16
<u>UTILIZATION AND CHARACTERISTICS OF SELECTED "OPTIONAL" ELEMENTS</u>	<u>C4</u>
CONTROLLED VOCABULARY-INDEX TERMS-CONTROLLED.....	C4.1
CONTROLLED VOCABULARY-THESAURUS.....	C4.2
LOCAL SUBJECT INDEX.....	C4.3
AVAILABILITY-RESOURCE DESCRIPTION	C4.4
METHODOLOGY	C4.5
RESOURCE PROFILE.....	4.3
RECORD TYPES	P1
OBJECTS REPRESENTED	P2
RECORD AGGREGATION.....	P3
CONTAINERS	P4
SERVICEABILITY	4.4
FILE FORMATS.....	S1
CONTENT HYPERTEXT.....	S2
CAPITALIZATION.....	S3
INDENTATION	S4
ELEMENT DISPLAY ORDER.....	S5
DEFINITIONS OF ACRONYMS	S6
CITATION OF LEGISLATION	S7
LOCALLY DEFINED ELEMENTS.....	S8

2.0. METHOD OVERVIEW

The analysis was performed in two phases: Phase 1 comprised examination of a pool of 80 records from 40 agencies' GILS retrieved deliberately to represent a range of information resource types (e.g., databases, catalogs, records systems). These records served as the basis for developing and operationalizing a set of more than 50 qualitative and quantitative evaluative criteria that included records' format, aggregation, media representation, and descriptiveness. Descriptiveness was defined as the incidence of utilization and content (value) attributes for all mandatory and selected optional elements and subelements specified by *FIPS Pub. 192 Annex E-GILS Core Elements* definitions as reproduced and supplemented by usage guidelines and examples in the *NARA Guidelines*. In Phase 2, these criteria were systematically applied to a set of 83 records randomly retrieved January 13 and 14, 1997, from 42 agencies' GILS. Results, therefore, reflect record content *at the time of retrieval* and represent a "snapshot" during only one, and arbitrary, point in the GILS system lifecycle. In addition, the "Core subset" analysis comprises records within the total sample that contained a value of "US Federal GILS" or "U.S. Federal GILS" in the Controlled Vocabulary-Local Subject Index Term subelement.

Appendix C-4 Record Content Analysis Methodology provides a complete discussion of methodology, and Appendix D-4 Record Content Analysis Instrument presents the database fields used for data collection.

3.0. DATA SUMMARIES

The following tables summarize significant data resulting from the analysis in terms of the objective-based categories outlined in Section 1.1 Organization of Material. Operational definitions of semantics used during the analysis are presented in Appendix C-4 Record Content Analysis Methodology and reiterated as applicable in Section 4.0 Findings, Discussion, and Recommendations. Investigators strongly recommend that interpretation of the following findings be guided by the complete data and discussions provided in Section 4.0.

Section 3.5 High Quality Records From the Sample cites records exemplifying many of the important quality characteristics. These records are reproduced (as printed directly from the Web) in Attachments E2-1a through E2-1d to this appendix.

3.1. Accuracy

The following summary data reflect the number of errors identified in the sampled GILS records. (The scope of the current study did not include verification of the "accuracy" of substantive information *per se*—such as telephone numbers, URLs, etc.)

Criteria	Data Summary and Highlights	Section 4.0
Records with incorrect date formats	33%	C3.16
Records with file formatting errors	24%	A1
Records with spelling or typographical errors	10%	A2

3.2 Completeness

This table presents summary data concerning the fullness of sampled records in terms of inclusion of both "mandatory" and "optional" elements (as defined by the NARA *Guidelines*).

Criteria	Data Summary and Highlights	Section 4.0
General		
Number of populated elements per record	max 190 min 11 avg 42	C1
Records containing "blank" (labeled but null value) elements	36%	C2
Locally defined elements identified	12	S8
<i>Utilization of 12 mandatory elements (excludes AIS-Mandatory and RM-dependent)</i>		
Total sample	96%	C3.1-16
Core subset	95%	C3.1-16
<i>Utilization of selected optional elements</i>		
Controlled Vocabulary	12% LCSH most popular thesaurus	C4.1 C4.2
Local Subject Index	54% (includes variants of "US Federal GILS")	C4.3
Availability-Resource Description	12% for total sample 24% for core subset	C4.4
Methodology	2%	C4.5
Originator subelement(s)	65%	C3.2
Record Source subelement(s)	63%	C3.15

3.3 Resource Profile

The following table summarizes characteristics of the resources described in the sampled records.

Criteria	Data Summary and Highlights	Section 4.0
Record types (AIS, Locator, Privacy Act systems)	could not be discerned	P1
Objects represented	(12 types) 22% "subject matter database" 19% "publication" 59% other	P2
Aggregation	(5 levels) 36% Record aggregated objects 25% Aggregated object represented 20% Discrete object 12% Object aggregates metadata 6% Unknown	P3
Containers	(7 types) 22% Multiple 23% Print 8% Web	P4

3.4. Serviceability

The "serviceability" data summarized below are considered to represent record effectiveness in terms of the degree to which they enhance NIDR, convenience to the user, aesthetics, readability, and relevance judgment.

Criteria	Data Summary and Highlights	Section 4.0
NIDR Factors		
Records with spelling or typographical errors	10%	A2
Records with Controlled Vocabulary	12% LCSH most popular thesaurus	C4.1 C4.2
Records with Local Subject Index	54% 6% of Core subset contained "U.S." rather than "US" in <US Federal GILS>	C4.3 C3.3
Records with Resource Description	12% for total sample 24% for core subset	C4.4
Records with Record Schedule number	14%	C3.13
Records with Control Identifier	91%	C3.14
Records with Originator subelement(s)	65%	C3.2
Records with Record Source subelement(s)	63%	C3.15
Locally defined elements identified	12	S9
Aggregation	36% Record aggregated objects 25% Aggregated object represented 20% Discrete object 12% Object aggregates metadata 6% Unknown	P3
Records with (any) acronyms	65%	S6
Records with undefined acronyms	12%	S6
Records with dates in incorrect format	33%	C3.16
Records with legislative citation	48% (only one was GILS-related)	S7
User Convenience		
Preferred element display order	64% for total sample 57% for Core subset	S5
Points of Contact	50% offices 23% personal names 9% job title 3% other	C3.12
Records with Availability-Distributor	93%	C3.7
Availability-Order Process	86%	C3.8
Records with hypertext	25% total 52% in Available Linkage 24% in Distributor Network Address	S2
File formats	81% ASCII 83% HTML 2% SGML 0% PDF	S1

(continued)

Criteria	Data Summary and Highlights	Section 4.0
Aesthetics/Readability		
Number of populated elements per record	max 190 min 11 avg 42	C1
Records containing "blank" (labeled but null value) elements	36%	C2
File formats	81% ASCII 83% HTML 2% SGML 0% PDF	S1
Records with file formatting errors	24%	A1
Capitalization style	86% sentence-case 10% elements-only capitalized	S3
Records employing any indentation pattern	73%	S4
Relevance-Judgment Factors		
Records with descriptive Titles	75%	C3.1
Records with descriptive Abstract	86%	C3.4
Records with (any) undefined acronyms	12%	S6
Records with substantive Access Constraints	29%	C3.10
Records with substantive Use Constraints	17%	C3.11
Records naming container (dissemination media)	50%	P4
Locally defined elements identified	12	S8
Records with Resource Description	12% for total sample 24% for core subset	C4.4
Records misusing Date of Last Modification as referring to resource rather than record	at least 4	C3.16

3.5. Examples of High-Quality Records from Sample

Four records exhibiting characteristics of "high quality" are provided as examples in Attachments E2-1a through E2-1d to this appendix.

- *AHCPR Publications Clearinghouse* available at <<http://www.dhhs.gov/progorg/oirm/newhhsgils.htm>> by searching Control Identifier (quotes required) "HHS-AHC-00509"
- *Aviation Accident Synopses World Wide Web Page* available by browsing <<http://www.ntsb.gov/Info/Info.htm>> or directly at <<http://www.ntsb.gov/Info/GILS/GILSSYN.htm>>
- *Farm Credit Administration's Privacy Act Systems* available <http://www.access.gpo.gov/su_docs/gils/gilsfld.html> by searching Control Identifier (quotes required) "FCA/PA-1"
- *FEMA Publications Catalog* available <http://www.access.gpo.gov/su_docs/gils/gils.html> by searching Federal Emergency Management Agency for "FEMA0001"

These records contain mandatory elements populated with NARA *Guidelines*-compliant values and are highly readable and descriptive without excessive length. In addition, they represent a range of "information objects" and "containers" (see Appendix C-4 Record Content Analysis Methodology): an information resource organization, an aggregated set of reports available via Web site, a "system of records" available via Government Printing Office (GPO) GPO Access, and a traditional printed publications catalog, respectively.

4.0. FINDINGS, DISCUSSION, AND RECOMMENDATIONS

Detailed results of the analysis are presented below in four major categories, which were assigned alphabetic codes. Section 4.1 "Accuracy (A)" concerns the incidence of errors. Section 4.2 "Completeness (C)" includes data concerning GILS record element utilization and values. Section 4.3 "Resource Profile (P)" includes general characteristics of the records studied. The final section, Section 4.4 "Serviceability (S)," presents findings relevant to record effectiveness in NIDR and user convenience. Further, each category code was coupled with a numeric code that reflects the order of data collection; under "Completeness" the data are additionally sorted in preferred element display order (per *FIPS Pub. 192* and the *NARA Guidelines*) for ease of reference.

4.1. Accuracy

The following data reflect the number of errors identified in the sampled GILS records. (The scope of the current study did not include verification of the "accuracy" of substantive information *per se*—such as telephone numbers, URLs, etc.) File formatting errors were found in about 25% of sampled records; spelling and typographical errors were evident in 10%.

A1 Formatting Errors

Findings: Roughly 1 in 4 records sampled contained error(s) attributable to file formatting or conversion, such as no hard-returns, unintentional hard-returns, incomplete files, HTML tags, stray ASCII text, file format characters, etc.

TOTAL SAMPLE

FILE FORMAT ERRORS	N	%
YES	20	24%
NO	63	76%
TOTAL	83	100%

CORE SUBSET

FILE FORMAT ERRORS	N	%
YES	7	17%
NO	35	83%
TOTAL	42	100%

Discussion: These data support record-creator complaints during focus groups and interviews with key informants that support personnel at times/places do not have online access with which to view GILS product. Study participants (see Appendix E-3 Scripted Online User Assessment Findings, Discussion, and Recommendations) showed poor tolerance of formatting errors.

Recommendations: Devise a hard-/software independent template and/or HTML editor for record formatting, or limit formatting responsibility to agency or subcontracted personnel with Web browsers.

A2 Spelling and Typographical Errors

Findings: The examination revealed that 1 in 10 records sampled contained spelling or typographical errors; the percentage was somewhat lower in the Core subset.

TOTAL SAMPLE

SPELLING ERRORS/TYPOS	N	%
YES	8	10%
NO	75	90%
TOTAL	83	100%

CORE SUBSET

SPELLING ERRORS/TYPOS	N	%
YES	3	7%
NO	39	93%
TOTAL	42	100%

Discussion: In addition to the possibility of record retrieval failure caused by lack of exact matching of user input, the presence of spelling and typographical errors may erode users' trust in the transmission and/or content integrity of records.

Recommendations: Use machine-based spell checkers, or assign checking responsibility to someone other than the writer.

4.2. Completeness

The following data concern the fullness of sampled records in terms of inclusion of both "mandatory" and "optional" elements (as defined by the NARA *Guidelines*). Of 67 possible element tags, the sample averaged 42 elements containing substantive values of a value of "none," "not applicable," etc. Nearly 40% of the records sampled featured the practice of presenting some "blank" elements—i.e., labels with no data.

C1 Number of Elements Per Record

Findings: In the total sample, records averaged a "length" of 42 elements, with a maximum of 190 and minimum of 11. The most frequent element count (mode) was 33, featured by 8% of the records. The Core subset's element utilization rates were not significantly different from the total sample.

TOTAL SAMPLE

TOTAL FIELDS USED		
SUM	3500	
AVG	42	
MAX	190	
MIN	11	
MODE	33	8%

CORE SUBSET

TOTAL FIELDS USED		
SUM	1873	
AVG	45	
MAX	190	
MIN	11	
MODE	33	10%

Discussion: *FIPS Pub. 192* currently specifies 67 tags. User perception of a record's "length" may be related to several factors, including C2-Practice of Presenting Blank Elements, S1-File Formats, S3-Capitalization, S4-Indentation, and S5-Element Display Order, as well as the number of elements utilized and repeated and the extent of their values. In the Scripted Online User Assessment, a record containing 14 (total) populated elements was judged "just right" in length by most users; one user felt that record space was generally "wasteful in relation to what you get" clarified as meaning "not actual documents." Of the 190 elements of the maximum-length record noted above, 170 were Point of Contact subelements in which 17 agency field stations were described (for a "records management program"). In addition, although this practice was not assessed systematically, the investigators noted several instances of pronounced redundancy in record content—e.g., one record contained "browser is required" or equivalent in four different elements.

Recommendations: Implement Z39.50-compliant clients which will enable the presentation of customized views of the record. To address the appropriateness of content and placement of values, a possible research project might isolate a random sample of selected data values and test whether GILS-cognizant vs. noncognizant users can place them in the "correct" (per published standards) elements. For those not trained in bibliographic control, and even those who are familiar with library cataloging processes, the NARA *Guidelines* are complex and at times ambiguous. Record creators as well as quality checkers may require a more straightforward, procedural, or specific instructions as well as a FAQ list, pocket guide reference, context-sensitive online help, etc. to select elements required to describe the resource and its availability appropriately.

C2 Practice of Presenting "Blank" Elements

Findings: More than one-third of the records examined presented one or more elements (labels) containing no data. No pattern was discernible across agencies as to the basis for this practice. However, it is possible that record creators are working with a "generic," inflexible in the interest of time economy or internal quality assurance initiatives.

TOTAL SAMPLE

"BLANK" ELEMENTS	N	%
YES	30	36%
NO	53	64%
TOTAL	83	100%

CORE SUBSET

"BLANK" ELEMENTS	N	%
YES	17	40%
NO	25	60%
TOTAL	42	100%

Discussion: See C1-Number of Elements Per Record for a note about user perceptions of record "length." In addition, users may perceive "blank" elements as agency negligence or system error.

Recommendation: A further analysis of how record creators are handling elements perceived as irrelevant or not necessary (i.e., the incidence of "none," "not applicable," "N/A," "not required" and similar null values vs. presentation of "blank" elements vs. omission of such elements altogether) may indicate that a procedural standard is in order. A useful adjunct to this research could assess the incidence of elements presented for "fielded searching" and agency rationale for selection.

C3 Utilization and Selected Characteristics of "Mandatory" Elements

Overall, utilization of GILS mandatory elements was very high. Excluding Agency Program, Sources of Data, and Schedule Number because the incidence of AIS as a record type could not be determined, the analysis revealed 96% utilization for the total sample and 96% for the Core subset. Given that 50% of the records featured some variant of "US Federal GILS" in the Local Subject Index, this close match may be interpreted to mean either a nearly ubiquitous appreciation of the intrinsic value of the mandatory elements or a similarly ubiquitous uncertainty as to the designation "core" record.

The characteristics of *values* found in mandatory elements was less positive, however:

- The incidence of descriptive titles was relatively low—only 75% for the total sample and 67% for the Core
- Descriptiveness of Abstracts was slightly better than that of Titles, at around 86%
- Substantive Access Constraints and Use Constraints were named in 29% and 17% of the records, respectively
- Almost 25% of records sampled named an individual as Point of Contact
- The granularity of organizational descriptions varied widely; Originator and Record Source subelement(s) were used in about 65% of records
- The analysis revealed that nearly 40% of the records contained at least one date in other than YYYYMMDD format.

These findings indicate that adherence to the NARA *Guidelines* in terms of element inclusion is high but that application of usage recommendations is less predictable.

The following results and discussion are presented per element in "preferred display" order for a GILS record (*FIPS Pub. 192* and NARA *Guidelines*). Note that the term utilization here means that the element was present, presented, and populated (even in cases where "none", "n/a", or similar acknowledgment of an absence of substantive content was present).

C3.1 Title

Findings: All records sampled utilized Title. Of the total sample, 3 in 4 records' Titles were coded descriptive; the incidence was slightly lower (67%) in the Core subset. The code of nondescriptive was applied to instances such as:

- Annual Reports failing to name the year
- a record titled "Employment Center," where the information object described was a telephonic job line
- use of the singular (e.g., "Report") when other element values implied more than one information object
- records titled simply "General Files" and "Minutes."

Examples of descriptive titles include "Investment Funds Brochure," "Automated Tariff Filing and Information System," and "GPO Access User Guide Online via GPO Access." In addition, although time constraints precluded a systematic assessment, the investigators noted at least the following terms used in titles of records describing an agency Web site: "Home Page," "Homepage," "World Wide Web Site," and "Internet Site."

TOTAL SAMPLE

TITLE DESCRIPTIVE	N	%
YES	62	75%
NO	21	25%
TOTAL	83	100%

CORE SUBSET

TITLE DESCRIPTIVE	N	%
YES	28	67%
NO	14	33%
TOTAL	42	100%

Discussion: The importance of descriptive Titles cannot be overemphasized given that it represents the only substantive content returned by most GILS search engines in the results list. The Scripted Online User Assessment determined that most users will judge a "hit's" relevancy first by appearance of their search terms in the Title, followed by WAIS "score," a finding related to a difficulty encountered during the current content analysis—particularly of titles of automated information systems. The NARA Guidelines state:

[1]This element conveys the most significant aspects of the referenced resource and is intended for initial presentation to users independently of other elements. It should provide sufficient information to allow users to make an initial decision on likely relevance. It should convey the most significant information available, including the general topic area, as well as a specific reference to the subject.

[2]The title provides the name of the information resource as assigned by the agency. For automated information systems and locators to information dissemination products, the title is the officially assigned name for the system.

This guideline assumes that the names of automated information systems (see P5-Object Represented) are descriptive for purposes of NIDR—an assumption the investigators find unwarranted in experience (the choice of name for a system is often motivated by its ability to create a facile acronym) and by way of this investigation. For example, are "OpenNet," "Enterprise Information System: EIS," "OEPC BBS," and "HUD USER," *descriptive* titles? According to [1] above, no. NARA, in their record "CLIO, The National Archives Information Server" attempts to accommodate both principles (and the investigators assume, although the term is not defined in the record, that "CLIO" spelled out might not be "descriptive."). The latter example brings to light another issue—that of search engines failure to return the agency acronym with the search result (title). This further degrades the NIDR value of titles such as "Consumer Bulletins," which emphasize audience at the expense of "the general topic area, as well as a specific reference to the subject" (per [2] above)—a problem that appears to be exacerbated upon wide cross-agency searching. Unfortunately, on the other side of this coin are titles such as "Federal Communications Commission (FCC) Technical Documents Created By Its Office of Engineering and Technology (OET)," a title that some might characterize as verbose. Section S2-Content Hypertext also discusses Title values.

Recommendations: Given that the title is the only record-content cue provided to current GILS users, it is essential that some degree of objective, third-party (i.e., other than resource creator and/or record creator) evaluation be applied when evaluating title descriptiveness. Implementation of Z39.50-compliant systems could eliminate this problem by allowing the user to select presentation of the Abstract with the Title to assist in judging relevancy.

C3.2 Originator

Findings: All records examined utilized the Originator element. 65% of the records sampled utilized at least one Originator subelement but only 45% presented its label.

TOTAL SAMPLE

ORIGINATOR SUBELEMENTS	N	%
YES	54	65%
NO	29	35%
TOTAL	83	100%

CORE SUBSET

ORIGINATOR SUBELEMENTS	N	%
YES	24	57%
NO	18	43%
TOTAL	42	100%

TOTAL SAMPLE

ORIGINATOR SUBELEMENTS LABELED	N	%
YES	37	45%
NO	46	55%
TOTAL	83	100%

CORE SUBSET

ORIGINATOR SUBELEMENTS LABELED	N	%
YES	16	38%
NO	26	62%
TOTAL	42	100%

Discussion: The NARA *Guidelines* provide the following definition of this element: "This element occurs once per locator record. It identifies the information resource originator, named as in the *U.S. Government Manual* where applicable." Most study participants felt that "all GILS records should look alike" (see Appendix E-3 Scripted Online User Assessment Findings, Discussion, and Recommendations). It may be concluded that this preference refers to the presence and absence of display characteristics rather than content. An assessment of use of agency names as set forth in the *U.S. Government Manual* was not performed due to time constraints, however it is noted that this requirement will not serve the stated purpose of supporting NIDR unless the user has ready access to the *Government Manual* Table of Contents or Appendix A: Commonly Used Abbreviations and Acronyms.

Recommendations: The term "information resource originator" is undefined and its relationship to Point of Contact's "organizational unit that *created and maintains* [investigators' emphasis] the information dissemination product or information system" is unclear. It is suggested that research be conducted to assess users' expectations concerning presentation of true "authorship" data in the bibliographic sense as opposed to the entity responsible for compilation, administrative maintenance, or dissemination of the resource. In addition, it is suggested that a cross-tabulation of Originator/Point-of-Contact/Record Source values be performed and the values sampled for accuracy to verify potential confusion among definitional terms and roles implied by "originator," "creator," "provider," etc. (See discussions at C3.12-Point of Contact and C.15-Record Source.) Further assessment of user preferences for display of organizational subelement labels and values, as well as their aggregation levels and resultant maintenance burden, is recommended as well. Investigators believe that standardization of element label display will contribute to users' "footing" within GILS vs. other information space—e.g., recognition that GILS is a bounded (by function), top-down, two-dimensional service that spans across all agencies. In addition, on GPO Access GILS, it may be prudent to provide a hypertext link from "*US Government Manual*" in the field definition files to a recast version of *Government Manual* Appendix A: Commonly Used Abbreviations and Acronyms.

C3.3 Local Subject Index—"US Federal GILS"

Findings: As noted at the beginning of this section, for purposes of the GILS Evaluation record-content analysis, records containing "US Federal GILS" or "U.S. Federal GILS" in this element constituted the analysis subset called "Core subset." The NARA-recommended "US" format appeared in 43% of the records sampled.

TOTAL SAMPLE

US FED GILS	N	%
US FEDERAL GILS	36	43%
U.S. FEDERAL GILS	5	6%
NOT USED	41	49%
"US" BUT IN WRONG ELEMENT	1	1%
TOTAL	83	100%

CORE SUBSET

US FED GILS	N	%
US FEDERAL GILS	36	86%
U.S. FEDERAL GILS	5	12%
NOT USED	0	0%
"US" BUT IN WRONG ELEMENT	1	2%
TOTAL	42	100%

Recommendations: Investigators recommend that the concept and functionality of the GILS "Core" be re-examined in light of the study's overall findings and given the lack of significant difference in utilization and quality of "mandatory" vs. "optional" elements revealed during this content analysis. Should the requirement remain viable, the results above call for a clarification of purpose and implementation guidelines in GILS standards and procedures.

C3.4 Abstract

Findings: Only 2 records in the sample failed to provide a value in this element, one of these was in the Core subset. More than 85% of sampled records Abstracts were coded "descriptive." "Nondescriptive" incidents included:

- "none."
- "The [agency] is responsible."
- Values equaling the name or title of the information resource
- Investigator judgment that content matched more closely another element's definition (e.g., Purpose or Availability)
- Investigator judgment that the content was degraded by use of incomplete sentences or technical jargon or "bureaucrat-ese" (e.g., "It discusses the application of one or more provisions of law to the detailed factual situation set forth as a proposed course of conduct in light of requirements of a particular...regulation rule [sic], order to cease and desist or affirmative order...").

TOTAL SAMPLE

ABSTRACT DESCRIPTIVE	N	%
YES	71	86%
NO	10	12%
NOT USED	2	2%
TOTAL	83	100%

CORE SUBSET

ABSTRACT DESCRIPTIVE	N	%
YES	37	88%
NO	4	10%
NOT USED	1	2%
TOTAL	42	100%

Discussion: Per the NARA *Guidelines*:

This [element's] narrative should provide enough general information to allow the user to determine if the information resource has sufficient potential to warrant contacting the provider for further information...The content of the abstract will be dependent upon the nature of the entity to be described (i.e., a locator to information dissemination products, a Privacy Act system, or an automated information system). The abstract may include, but is not limited to, discussion of the information content (including data coverage, persons, events, and topics); form of information; media; time span; and geographic coverage.

The Abstract is a familiar and preferred basis of relevancy judgment for many users. Creation of descriptive abstracts is time-consuming and especially difficult for personnel unfamiliar with the resource subject and/or abstracting principles.

Recommendations: More specific guidelines, and perhaps exposure to a greater number of effective (model) Abstracts for various information objects (see P5-Information Object), may assist record creators in developing consistency in the descriptions. In addition, investigators recommend that the definition, usage, and structure of the Resource Description subelement be revised to provide physical description of the object as recognizable by the user

rather than by the distributor, that this information be removed from the Abstract element to the Resource Description subelement, and that the subelement be mandatory and structurally associated with the Abstract element. See C4.4 Availability-Resource Description.

C3.5 Purpose

Findings: Of the total sample, 96% of records utilized this element; of the Core subset, 95% did.

TOTAL SAMPLE

PURPOSE	N	%
YES	80	96%
NO	3	4%
TOTAL	83	100%

CORE SUBSET

PURPOSE	N	%
YES	40	95%
NO	2	5%
TOTAL	42	100%

Discussion: The NARA *Guidelines* state that this element "describes why the information resource is offered and identifies other programs, projects, and legislative actions wholly or partially responsible for the establishment or continued delivery of this information resource." S7-Citation of Legislation presents data concerning the total (element-wide) incidence; the NARA Guidelines also call for legislative references in the Agency Program element (see C3.6-Agency Program).

Recommendations: The feasibility of automating the insertion of substantively correct and properly formatted legislative citations upon inclusion of a program or project name (i.e., installation of an expert-system legislation index cross-reference macro) could reduce record creation and maintenance burden as well as facilitate identification of legislation-dependent resources for agency users (IRM and public information office personnel), Congressional users, and public policy researchers.

C3.6 Agency Program (mandatory if AIS)

Findings: The following tables present utilization of this element across the entire sample population because "AIS" as a record type could not be determined (see P4-Record Type). More than 70% of sampled records utilized this element. While time constraints precluded a systematic analysis of Agency Program values, the investigators noted only one record's reference to "OMB Circular A-130."

TOTAL SAMPLE

AGENCY PROGRAM (AIS)	N	%
YES	63	76%
NO	20	24%
TOTAL	83	100%

CORE SUBSET

AGENCY PROGRAM (AIS)	N	%
YES	33	79%
NO	9	21%
TOTAL	42	100%

Discussion: See P5-Information Object; given a liberal interpretation that AISs comprise "subject-matter databases," "agency homepages," "bibliographic databases," and "systems of systems," one may extrapolate only a predicted 50% utilization of this element. It is possible that "non-AIS resource" record creators are recognizing the intrinsic value of this element in educating public users about agency functions. NARA *Guidelines* state that "This element identifies the major agency program or mission supported by the system and should include a citation for any specific legislative authorities associated with this information resource...In general terms, it explains why the information resource was created in the first place. The rationale for a specific design is found in the PURPOSE element."

Recommendations: The rationale for isolating mandatory use of this element to describe AISs as well as differentiation of expected values between Purpose and Agency Program should be clarified in the NARA *Guidelines*. See also C3.5-Purpose recommendation concerning legislative citation.

C3.7 Availability-Distributor

Findings: 90% of records examined utilized at least one subordinate field of the Availability- Distributor subelement. Incidence of use in the Core subset was identical to that of the total sample.

TOTAL SAMPLE

AVAILABILITY- DISTRIBUTOR	N	%
YES	77	93%
NO	6	7%
TOTAL	83	100%

CORE SUBSET

AVAILABILITY- DISTRIBUTOR	N	%
YES	39	93%
NO	3	7%
TOTAL	42	100%

Discussion: Per the NARA *Guidelines*, the mandatory Availability element “is a grouping of subelements that together describe how the information resource is made available.” Instructions for the mandatory Distributor subelement state: “Complete as many of the subordinate fields as necessary to identify the party from whom the information resource is available.” Subordinate fields are Distributor’s Name, Organization, Street Address, City, State, Zip Code, Country, Network Address, Hours of Service, Telephone, and Fax. The word choice “necessary to identify to the party” may be too vague to promote standard usage.

Recommendations: The *Guidelines* wording “Complete as many of the subordinate fields as necessary to identify the party” could be improved to “Complete all subordinate fields as available to assist the user in communicating with the Distributor.”

C3.8 Availability-Order Process

Findings: Utilization (86%) of at least one subordinate field of the Availability- Distributor subelement was identical for the total sample and Core subset. One agency’s sampled records contained a request for users to specify that the order is “in response to information in a...GILS record.”

TOTAL SAMPLE

AVAILABILITY-ORDER PROCESS	N	%
YES	71	86%
NO	12	14%
TOTAL	83	100%

CORE SUBSET

AVAILABILITY-ORDER PROCESS	N	%
YES	36	86%
NO	6	14%
TOTAL	42	100%

Discussion: Per the NARA *Guidelines*, the mandatory Availability element “is a grouping of subelements that together describe how the information resource is made available.” Instructions for the mandatory Order Process subelement state:

Provide information on the common ways in which copies of the information resource or data from the resource may be obtained from this distributor. Alternatively, provide information on how the resource may be accessed. Note here if there is an electronic linkage for ordering, and also complete the LINKAGE subelement. Each agency may wish to establish standard ordering instructions for inclusion here.

The 86% utilization rate, compared with those of other mandatory elements and subelements, may be relatively low as a result of record creators entering “ordering” information in other elements [e.g., Access Constraints (see C3.10) or Linkage as noted above] or omission of this element altogether for nonaccessible resources (e.g., proprietary databases) or those that can not be “ordered” per se, such as an agency library.

Recommendations: Content analysis of the *Guidelines* by a third-party (i.e., not someone who was involved with the writing) is recommended to reveal areas of redundancy and ambiguity. Another approach would involve isolation of a random sample of selected data values (e.g., a distributor’s URL) and testing whether GILS-cognizant vs. noncognizant users can place them in the “correct” (per published standards) elements.

C3.9 Sources of Data (mandatory if AIS)

Findings: The following tables present utilization of this element across the entire sample population because "AIS" as a record type could not be determined (see P4-Record Type). More than 50% of the records in the total sample utilized this element; utilization was lower (43%) in the Core subset.

TOTAL SAMPLE

SOURCE OF DATA (AIS)	N	%
YES	42	51%
NO	41	49%
TOTAL	83	100%

CORE SUBSET

SOURCE OF DATA (AIS)	N	%
YES	18	43%
NO	24	57%
TOTAL	42	100%

Discussion: See P5-Information Object; given a liberal interpretation that AISs comprise "subject-matter databases," "agency homepages," "bibliographic databases," and "systems of systems," one may extrapolate the found 50% utilization of this element. However, although values of this element were not analyzed systematically, the investigators noted several instances of non-AIS usage (e.g., the record "Detailed Briefing Materials," which describes a print supplement to the agency's Annual Budget press release, cites the office responsible). NARA *Guidelines* instruct: "Give information about the primary sources or providers of data to the system. State if the information in the system is generated by the agency, or if it is received by the system from outside the agency. Examples of the source of information for a system from outside the agency include corporations doing business in the U.S., broadcast license holders, or another Federal agency. This practice possibly indicates that record creators recognize the intrinsic value of this element in educating public users about agency methods and procedures for collecting information."

Recommendations: The rationale for isolating mandatory use of this element to describe AISs should be clarified in the NARA *Guidelines*.

C3.10 Access Constraints

Finding: This element was utilized in 99% of the total sample's records and in 100% of Core subset's records. The tables below reveal a slight difference between samples' incidence of "none" (the NARA-recommended null value expression) vs. substantive values. More than 25% of the sampled records describe a resource with at least one access constraint.

TOTAL SAMPLE

ACCESS CONSTRAINT	N	%
"NONE"	58	70%
SUBSTANTIVE	24	29%
NOT USED	1	1%
TOTAL	83	100%

CORE SUBSET

ACCESS CONSTRAINT	N	%
"NONE"	33	79%
SUBSTANTIVE	9	21%
NOT USED	0	0%
TOTAL	42	100%

Discussion: See C3.8-Order Process concerning possible redundancy of NARA-recommended Access Constraint values; *Guidelines* instructions for Access Constraint state to "[include] other special restrictions or limitations on obtaining the information resource. Guidance on obtaining any users' manuals or other aids needed for the public to reasonably access the information resource must also be included here [in the Documentation subelement]." (The latter direction raises an additional issue of whether recommended "users' manuals or other aids" are to be described by separate but complementary GILS record(s) and, if so, whether the at-hand GILS record's Cross-Reference value or Documentation value should link to the actual manual or to the GILS record describing it. GPO's GILS records entitled "GPO Access" (Control Identifier: "GPO/SOD/OEIDS00027" and "GPO Access User Guide Online via GPO Access" (Control Identifier: "GPO/SOD/OEIDS-00026")) provide an example of this scenario. In addition, see discussion at C3.11-Use Constraints for possible confusion over "access" vs. "use" constraint elements' values.

Recommendations: Investigators recommend that the NARA *Guidelines* differentiate placement of certain constraints (e.g., "requires forms-capable browser"—in Access Constraint vs. Order Process) and clarify the relationship between or the presence/absence of a user's manual or aid and "access constraints."

C3.11 Use Constraints

Findings: This element was utilized in 100% of both the total sample's and Core subset's records. The tables below reveal a slight difference between samples' incidence of "none" (the NARA-recommended null value expression) vs. substantive values.

TOTAL SAMPLE

USE CONSTRAINT	N	%
"NONE"	69	83%
NAMED	14	17%
TOTAL	83	100%

CORE SUBSET

USE CONSTRAINT	N	%
"NONE"	34	81%
NAMED	8	19%
TOTAL	42	100%

Discussion: While analysis of the substantive values in this element was not within the scope of the current study, the investigators noted several instances of Use Constraint values referring to Access Constraints (e.g., "[agency service organization] does not lend reference materials...although the public is welcome to use all of these in-house"). One record populated both the access and use constraint elements with "This system is an internal information and processing system and is not generally available for review outside the agency." Use Constraints is defined by the NARA *Guidelines* as: "describ[ing] any constraints or legal prerequisites for using the information resource or its component products or services." The definition of Access Constraints (see C3.10 above) carries precisely the same wording save for the substitution of the word "accessing" for "using." However, NARA's usage guidelines, which contain differentiating caveats (e.g., "Restrictions on what may be done with the information once it has been accessed are found in the USE CONSTRAINTS data element"), for both these elements appear to have been missed by some record creators.

Recommendation: Upon confirmation that confusion among Access vs. Use Constraints is significant, it is recommended that Use Constraints be renamed "Restrictions on Use of Information" or some other more straightforward phrase.

C3.12 Point of Contact

Findings: Assessment of utilization for this element considered a substantive value in any subelement. 99% of the total sample and 98% of the Core subset records utilized this element. "Point of Contact Type" below reflects the initial subelement value in this element. 1 in 4 records sampled featured a personal name; roughly half featured the name of an office. 20% of the values were a job title.

TOTAL SAMPLE

POINT OF CONTACT TYPE	N	%
NAME	19	23%
OFFICE	45	54%
JOB TITLE	16	19%
OTHER	2	2%
NOT USED	1	1%
TOTAL	83	100%

CORE SUBSET

POINT OF CONTACT TYPE	N	%
NAME	7	17%
OFFICE	24	57%
JOB TITLE	8	19%
OTHER	2	5%
NOT USED	1	2%
TOTAL	42	100%

Discussion: Per the NARA *Guidelines*, this element "identifies an organization, and a person where appropriate, serving as the point of contact plus methods that may be used to make contact. This element consists of the following subelements: [Name; Organization; Street Address; City; State; Zip Code; Country; Network Address; Hours of Service; Telephone; Fax]...Complete as many of the subordinate fields as are necessary to identify the organization and individual *responsible for the content of the information dissemination product or automated information system* [investigators' emphasis]. While confirming the accuracy of Points of Contact was beyond the scope of the current analysis, it is doubtful that "authors" are being listed in this element in cases where discrete information products (see P5-Objects Represented) are being described.

Recommendation: It is suggested that research be conducted to assess users' expectations concerning presentation of true "authorship" data in the bibliographic sense as opposed to the entity responsible for compilation, administrative maintenance, or dissemination of the resource—a value they may expect feasibly to find in Originator (whose definition includes "organizational unit that *created and maintains* [investigators' emphasis] the information dissemination product or information system" and Record Source elements as well. In addition, continuous and unscheduled audit of the accuracy of Points of Contact, as well as the quality of Contacts' responses to GILS record-related questions, is essential. Users may not only abandon GILS as a result of a discrepancy in this element but broadcast the failure among communities of interest.

C3.13 Schedule Number (mandatory if intended to meet the obligation...to inventory automated information systems or other records series for records management purposes)

Findings: The following tables present incidence of populated Schedule Number elements and their respective values. For the total sample, 55% of records utilized this element. 14% of all records sampled contained a record schedule number; 41% contained values indicating that scheduling was not required, is pending, etc. For the Core subset, an identical 55% element utilization rate was found. 5% of Core records contained a schedule number and 50% contained values indicating that scheduling was not required, is pending, etc.

TOTAL SAMPLE

SCHEDULE NUMBER	N	%
NOT USED (TOTAL)	37	45%
"NOT SCHEDULED"	16	
"N/A"	4	
"PENDING"	3	
"NONE"	3	
"NOT APPLICABLE"	2	
"UNSCHEDULED"	2	
"SCHEDULE IN PROGRESS"	2	
THIS IS A NONRECORD INFORMATION DISSEMINATION PRODUCT	1	
SCHEDULE TO BE SUBMITTED PENDING NARA INFORMAL REVIEW OF . . . PUBLICATIONS	1	
TOTAL	34	41%
ATF RCS 201, ITEM 140	1	
GENERAL RECORDS SCHEDULE NUMBER #20.9	1	
GRS 14 Sec. 6 and GRS 20 Sec. 9	1	
N1-138-88-2	1	
N1-266-77-2-92	1	
N1-309-87-002	1	
N1-522-95-1	1	
N1-95-88-2/62-9.11	1	
NC1-122-79-1, ITEM 11	1	
PBGC ITEM #67	1	
SCHEDULED- N1420-93-1, #26	1	
SCHEDULED-N1-420-93-1, #12	1	
TOTAL	12	14%
GRAND TOTAL	83	100%

CORE SUBSET

SCHEDULE NUMBER	N	%
NOT USED (TOTAL)	19	45%
"NOT SCHEDULED"	11	
"N/A"	0	
"PENDING"	3	
"NONE"	3	
"NOT APPLICABLE"	2	
"UNSCHEDULED"	0	
"SCHEDULE IN PROGRESS"	2	
THIS IS A NONRECORD INFORMATION DISSEMINATION PRODUCT	0	
SCHEDULE TO BE SUBMITTED PENDING NARA INFORMAL REVIEW OF . . . PUBLICATIONS	0	
TOTAL	21	50%
N1-266-77-2-92	1	
N1-95-88-2/62-9.11	1	
TOTAL	2	5%
GRAND TOTAL	42	100%

Discussion and Recommendations: See Chapter 4 for discussion of GILS and records management.

C3.14 Control Identifier

Findings: This element was utilized in the total sample and Core subsets at 91% and 88%, respectively.

TOTAL SAMPLE

CONTROL ID	N	%
YES	76	91%
NO	7	9%
TOTAL	83	100%

CORE SUBSET

CONTROL ID	N	%
YES	37	88%
NO	5	2%
TOTAL	42	100%

Discussion: NARA *Guidelines* state that "this element is defined by the *information provider* [investigators' emphasis] and is used to distinguish this *locator record* [investigators' emphasis] from all other GILS Core entries. The control identifier should be distinguished with the record source agency acronym as provided in the *U.S. Government Manual*....Create a unique identifying number for each GILS Core entry. The control identifier will consist of two parts: an identifying acronym followed by a control number." While no explanation of the lower utilization rate for the Core subset can be offered by the investigators, evidence was found of two possibly related problems.

First, college-educated participants in the online user assessment (see Appendix E-3 Scripted Online User Assessment Findings, Discussion, and Recommendations) did not understand the definition or use of this element. Upon presentation of the term only, prior to any searching, 8 of 10 stated they were "uncomfortable with, or unsure of using" the element; upon presentation of the NARA definition, 4 of 8 respondents stated that "the definition, and how it fits into GILS" was still "unclear." It may not be unreasonable to assume that the users' experience is transferable to a record-creator.

Second, "information provider" is an undefined but crucial term within the element's definition and could cause confusion when opposed to the construct of "author" or distributor (see further discussion of this issue at C3.12-Point of Contact). In addition, the definition assumes an understanding of the concept of "core" and "locator" records that may be unwarranted even among those in the GILS community—given the overall utilization rate of mandatory elements by the Core subset compared with the total sample. Finally, the *Guidelines* instructions concerning the values for this element *vis-a-vis* the many if/then scenarios are relatively elaborate and may require a level of inter-agency cooperation not in evidence during this study.

An assessment of use of agency acronyms as set forth in the *U.S. Government Manual* was not performed due to time constraints, however it is noted that this requirement will not serve the stated purpose of supporting NIDR unless the user has ready access to the *Government Manual*.

Recommendations: Further assessments of utilization rates, including an accounting for the quality of values as well as the rationale used in their construction/assignment appears warranted. In addition, on GPO Access GILS, it may be prudent to provide a hypertext link from "*US Government Manual*" in the field definition files to a recast version of *Government Manual* Appendix A: Commonly Used Abbreviations and Acronyms.

C3.15 Record Source

Findings: 90% of the total sample's records utilized this element; 88% of the Core subset records did. 63% of the records sampled utilized at least one Record Source subelement but only 46% presented its label.

TOTAL SAMPLE

RECORD SOURCE	N	%
YES	74	90%
NO	9	10%
TOTAL	83	100%

CORE SUBSET

RECORD SOURCE	N	%
YES	37	88%
NO	5	2%
TOTAL	42	100%

TOTAL SAMPLE

RECORD SOURCE SUBELEMENTS	N	%
YES	52	63%
NO	31	37%
TOTAL	83	100%

CORE SUBSET

RECORD SOURCE SUBELEMENTS	N	%
YES	27	64%
NO	15	36%
TOTAL	42	100%

TOTAL SAMPLE

RECORD SOURCE SUBELEMENTS LABELED	N	%
YES	38	46%
NO	45	54%
TOTAL	83	100%

CORE SUBSET

RECORD SOURCE SUBELEMENTS LABELED	N	%
YES	16	38%
NO	26	62%
TOTAL	42	100%

Discussion: "This element identifies the organization, as named in the *U.S. Government Manual*, that created or last modified this locator record...Give the name of an organization, and normally the name of the unit, that has created this GILS Core entry" per the NARA *Guidelines*. Most study participants felt that "all GILS records should look alike" (see Appendix E-3 Scripted Online User Assessment Findings, Discussion, and Recommendations), and it may be concluded that this preference refers to the presence and absence of display characteristics rather than content. Investigators believe that standardization of element label display will contribute to users' "footing" within GILS vs. other information space—e.g., recognition that GILS is a bounded (by function), top-down, two-dimensional service that spans across all agencies. An assessment of use of agency names as set forth in the *U.S. Government Manual* was not performed due to time constraints, however it is noted that this requirement will not serve the stated purpose of supporting NIDR unless the user has ready access to the *Government Manual* Table of Contents or Appendix A: Commonly Used Abbreviations and Acronyms.

Recommendation: Further assessment of user preferences for display of organizational subelement labels and values, as well as their aggregation levels and resultant maintenance burden, is recommended. In addition, on GPO Access GILS, it may be prudent to provide a hypertext link from "US Government Manual" in the field definition files to a recast version of *Government Manual* Appendix A: Commonly Used Abbreviations and Acronyms.

C3.16 Date of Last Modification

Findings: Utilization of this element was identical in the total and Core subset samples: 98%. The tables below also reveal a high +33% incidence of nonstandard date formats in sampled GILS records (where N=1 instance of nonstandard date, in *any* element). In addition, the data reveal that at least four sampled records contained values indicating a misperception that the value of the element refers to the information resource being described rather than the record itself: "FY 1994," "1988," "1989," and "continuously updated."

TOTAL SAMPLE

DATE OF LAST MOD	N	%
YES	81	98%
NO	2	2%
TOTAL	83	100%

CORE SUBSET

DATE OF LAST MOD	N	%
YES	41	98%
NO	1	2%
TOTAL	42	100%

TOTAL SAMPLE

LAST MOD DATE	
NOT USED (N)	2
MOST RECENT	THURSDAY, DECEMBER 12, 1996
EARLIEST	SATURDAY, DECEMBER 31, 1988
MODE	TUESDAY, DECEMBER 12, 1995
AVERAGE	SUNDAY, DECEMBER 31, 1995

CORE SUBSET

LAST MOD DATE	
NOT USED (N)	2
MOST RECENT	THURSDAY, DECEMBER 12, 1996
EARLIEST	MONDAY, DECEMBER 12, 1994
MODE	TUESDAY, DECEMBER 12, 1995
AVERAGE	SATURDAY, MARCH 16, 1996

TOTAL SAMPLE

ALL DATES IN YYYYMMDD?	N	%
YES	51	61%
NO	32	39%
TOTAL	83	100%

CORE SUBSET

ALL DATES IN YYYYMMDD?	N	%
YES	25	60%
NO	17	40%
TOTAL	42	100%

Discussion: The NARA *Guidelines* definition of this element is: "This element identifies the latest date on which this locator record was created or modified"—a relatively straightforward statement that is unfortunately obfuscated by the following [investigators' emphasis]:

Usage Guidelines: This element is used initially to record the date of the creation of the GILS Core *entry*. The value of the element changes with any subsequent modifications [*to what?*]. Only the date of the last modification of the *entry* needs be included in the GILS Core *entry*, but agencies may wish to track the date of the initial creation of the GILS *record* as well in a local [sic] defined element.

An interview with GPO Access GILS support staff corroborated record-creators' confusion concerning whether the value of this element describes the information resource or the GILS record itself. In addition, more than half of participants in the scripted online user assessment chose the value appearing in this element over others when asked to supply "how up-to-date are the described materials [in the record]?"

Recommendations: It is highly recommended that the Date of Last Modification element be renamed "Record Revision Date" and the confusing "entry" nomenclature in the usage guidelines be eliminated. In addition, where software resources permit, "auto-correct/auto-format" macros should be utilized to standardize dates. Use of computer-generated time/stamp dates, perhaps as a final "record checked and approved for release" procedure, may promote responsibility for record content quality among creators, facilitate audit sampling, and ease record maintenance as well. A possible future research project could assess government-wide frequency of change and clustering of Dates of Last Modification to reveal patterns in record maintenance.

C4 Utilization and Characteristics of Selected "Optional" Elements

The following list summarizes findings from this portion of the analysis.

- Utilization of Controlled Vocabulary was assessed at 10%, with *Library of Congress Subject Headings* being the most often cited Thesaurus
- About half the records sampled featured at least one Local Subject Index term—nearly all being a variant of “US Federal GILS”
- Resource Description was provided in only 12% of the total sample but 24% of the Core subset
- Utilization of Methodology was negligible.

The relatively low incidence of nonmandatory elements in the sampled records is not seen as indicative of their potential application—i.e., nothing about the nature of the information resources described was revealed as precluding or discouraging their use. Rather it may reflect a minimum-compliance or expediency-based approach to record creation, the degree of comprehensibility of or intellectual accessibility to the NARA *Guidelines*, and/or insufficient training as to the value of elements in networked information discovery and retrieval (NIDR).

The following results and discussion are presented per element in “preferred display” order for a GILS record (*FIPS Pub. 192* and NARA *Guidelines*). Note that the term utilization here means that the element was present, presented, and populated (even in cases where “none”, “n/a”, or similar acknowledgment of an absence of substantive content was present).

C4.1 Controlled Vocabulary-Index Terms-Controlled

Findings: 10% percent of sampled records utilized controlled index terms; for the Core subset, utilization was slightly higher.

TOTAL SAMPLE

INDEX TERMS-CONTROLLED	N	%
YES	8	10%
NO	75	90%
TOTAL	83	100%

CORE SUBSET

INDEX TERMS-CONTROLLED	N	%
YES	5	12%
NO	37	88%
TOTAL	42	100%

Discussion: The NARA *Guidelines* state the benefits of this element very well: “One method of identifying possible GILS entries of interest will be provided through the ability in GILS to search the entire text of the entry, including the narrative description in the ABSTRACT element. More precise search results can be achieved through the use of the CONTROLLED VOCABULARY element.” There is no doubt that inclusion of controlled vocabulary terms greatly enhances NIDR. In addition, use of registered thesauri (see C4.2-Controlled Vocabulary-Thesaurus) may orient users in the resource subject domain; the majority of participants in the online assessment of GILS believed “It would be easier to search GILS records if they were grouped hierarchically by subject.” (See Appendix E-3 Scripted Online User Assessment Findings, Discussion, and Recommendations).

Recommendations: It is recommended that agencies pursue research into the effects of “familiar” and specialized Controlled Vocabularies on NIDR, in concert with the Library of Congress and GPO’s Cataloging Branch, to lessen users’ dependency on knowledge of agency mission and to increase precision of information retrieval. In addition, the value of providing thesaurus hyperlinkages warrants systematic study as does user preferences for “catalog-based” NIDR (e.g., the approach of Yahoo! and the Argus Clearinghouse). Users who find relevant records easily will probably focus less on GILS “user-ugliness” and serve as the most effective mode of promotion.

C4.2 Controlled Vocabulary-Thesaurus

Findings: Utilization of this element was higher for the Core subset (17%). The total sample named a thesaurus 8% of the time. *Library of Congress Subject Headings* was the most frequent value.

TOTAL SAMPLE

THESAURI	N	%
HAZARDOUS WASTER SUPERFUND	2	2%
DATABASE		
LCSH	4	5%
DTIC	1	1%
TOTAL	7	8%

CORE SUBSET

THESAURI	N	%
HAZARDOUS WASTER SUPERFUND	2	5%
DATABASE		
LCSH	4	10%
DTIC	1	2%
TOTAL	7	17%

Discussion and Recommendations: See C4.1-Controlled Vocabulary-Index Terms-Controlled.

C4.3 Local Subject Index

Findings: This element was used in 98% of the core subset (see note at data), and in more than half of the total sample.

TOTAL SAMPLE

LOCAL SUBJECT TERMS	N	%
YES	45	54%
NO	38	46%
TOTAL	83	100%

CORE SUBSET

LOCAL SUBJECT TERMS	N	%
YES	41	98%
NO*	1	2%
TOTAL	42	100%

*This results from the appearance of "US Federal GILS" in an element other than Local Subject Index.

Discussion: The utilization rate of this element within the Core subset is largely attributable to the minimal inclusion of the required "US Federal GILS" or variant. The NARA *Guidelines* define usage of the element for: "supplement[ing] an existing thesauri or in the absence of an acceptable listing...Identify significant subjects that apply to the information resource including broad concepts and unusual aspects of the system or product. Use topical subject headings consisting of general subject terms or names of events or objects."

Recommendation: An assessment of the frequency of use of Local Subject Terms other than to denote a "core" record is highly recommended as a first step for clarifying the usefulness of this element in public-access NIDR. It is possible that this element will evolve to describe "aspects of the system or product" and/or "names of events or objects" relevant to Internet information space navigation, for example: "gopher archive, listserv, SIG" or "census, PDF, Web download"; see also C4.4-Availability-Resource Description below.

C4.4 Availability-Resource Description

Findings: Only 12% the total sample's records utilized this element; utilization was double in the Core subset.

TOTAL SAMPLE

AVAILABILITY-RESOURCE DESCRIPTION	N	%
NO	73	88%
YES	10	12%
TOTAL	83	100%

CORE SUBSET

AVAILABILITY-RESOURCE DESCRIPTION	N	%
NO	32	76%
YES	10	24%
TOTAL	42	100%

Discussion: The NARA *Guidelines* read in part: "This subelement identifies the resource as it is known to the distributor." Difficulties encountered while characterizing GILS information resources (see P4 Record Types and P5-Objects Represented), aggregation levels (see P6-Aggregation), and dissemination media (see P7-Containers)

indicate that this element (or an additional element with this name) might better serve the objective of public-access NIDR if it were redefined to comprise the object/aggregation/container concept.

Recommendations: A qualitative analysis of element values for a large sample of GILS records is highly recommended. (See P4-Record Types, P5-Objects Represented, P6-Aggregation, and P7-Containers.)

C4.5 Methodology

Findings: Utilization of this element was negligible (2%) in the total sample; the Core subset reflected 5% utilization. One record's value for this element contained information about the manufacturing process for a CD-ROM (as opposed to the definitional content that would have described how the data on it were collected/compiled).

TOTAL SAMPLE

METHODOLOGY	N	%
NO	81	98%
YES	2	2%
TOTAL	83	100%

CORE SUBSET

METHODOLOGY	N	%
NO	40	95%
YES	2	5%
TOTAL	42	100%

Discussion: These findings may indicate another instance confusing application instructions. The NARA *Guidelines* state that the Methodology element "identifies any specialized tools, techniques, or methodology used to produce [investigators' emphasis] this information resource...Provide here information concerning significant methodological characteristics of the information resource. Examples of items that might be discussed include the algorithm, universe description, sampling procedures, classification, or validation characteristics." An incomplete reading of instructions might lead to the CD-ROM example cited above, and a record-creator's lack of knowledge of research terminology or access to relevant information might prohibit inclusion of a methodology description altogether.

Recommendations: This element, like Sources of Data (see C3.9) for AISs in particular, demands the contribution of resource creators or collaborators. Because its absence may subject the information resource to misuse or abandonment, it is recommended that agencies develop training and complementary procedures for record creators to recognize the need for and to obtain the information for populating this element.

4.3. Resource Profile

Appendix C-4 Record Content Analysis Methodology enumerates the 42 agencies whose records were included in the record content analysis. The following paragraphs provide chief defining characteristics of the sample.

Subject-matter databases (e.g., that for red cockaded woodpecker or accident-investigation information) comprised more than 20% of the resources described, followed by discrete publications (19%) and miscellaneous documents in an ad hoc collection (17%) (see P5-Objects Represented and Appendix C-4 Record Content Analysis Methodology for semantics). Agency homepages comprised 10% of the sampled records, a positive indicator of NIDR synergy among Federal information spaces. An analysis of aggregation types (see P6-Record Aggregation and Appendix C-4 Record Content Analysis Methodology for semantics) revealed that more than one-third of records described "new collections"—i.e., resources aggregated by virtue of the record itself (although the serviceability of this phenomenon in NIDR requires additional study). Related to description of resource types and aggregation, dissemination media or "containers" of information were found to be largely unidentifiable by the record's content; 22% of resources were described as being packaged/disseminated by multiple modes, and the highest identifiable single mode was print (23%) followed by Web (8%).

The most important finding of this "profile" analysis is that GILS covers a potentially unfathomable scope of information, and that the implications of this for record creation and NIDR demand consideration and commentary by policymakers. The resolution of issues such as:

- Optimal aggregation (i.e., unit of resource selected for representation)
- Enumeration and description of an increasing multiplicity of containers (and the potential for resultant resource derivation, amendment, or abridgment and ensuing authority, integrity, and maintenance concerns)

- Determination of how GILS should be presented to maximize users orientation in Federal information space *vis-a-vis* cross-agency search capabilities and synergy with agencies other metadata and full-text resources)

are perceived as fundamental to GILS ultimate success.

P1 Record Types

Findings: The investigators was unable to code sampled records according to record types of Automated Information System (AIS), Locator, and Privacy Act systems as outlined in the NARA *Guidelines*.

Discussion: (See also P5-Objects Represented.) The NARA *Guidelines* delineates three types of information resources: locators, automated information systems (AISs), and Privacy Act systems of records. The following text is provided to highlight the burden of denominating record types.

The *Guidelines* define "locators" by means of example:

...locators (not to be confused with the GILS itself) catalog or describe information products (such as books, CD-ROMs, publications, studies, reports, and patents, regardless of medium). For example, there could be a GILS Core entry describing the Library of Congress Information System (LOCIS). LOCIS consists of catalog entries for publications, and hence, serves as a locator to these publications

and "AISs" by means of reference to OMB Circular No. A-130:

An automated information system is a discrete set of information resources organized using information technology as defined in OMB Circular No. A-130

which, in turn, provides no definition of an AIS *per se* but defines "information resource" by way of inclusion:

The term "information resource" includes both government information and information technology

where:

The term "Information technology" means the hardware and software operated by a Federal agency...to accomplish a Federal function

as well as by functionality:

[AISs] may be used for the collection, processing, maintenance *or* [investigators' emphasis] dissemination of information, including Federal records...

A "Privacy Act system of records," the third information type delineated in the NARA *Guidelines*, is therein defined by way of reference to U.S.C. 552a, which specifies:

the term 'system of records' means a group of any records under the control of any agency from which information is retrieved by...[an] identifying particular assigned to the individual.

This brief definitional synopsis is presented to stress the difficulty, and perhaps futility, of classifying GILS records by the information types set forth in the NARA *Guidelines*. The definitions are not mutually exclusive (i.e., many "locators" and "Privacy Act systems" are also "AISs") and the investigators felt that intercoder reliability was sufficiently poor to abandon the task. However, it is noted that five records sampled specified "Privacy Act system."

The researchers found that a code of "object represented," developed during the current record content analysis, proved both more straightforward and productive in characterizing GILS records' content. Results of this approach are discussed in the following section.

P2 Objects Represented

Findings: This analysis aimed at capturing what types of information products and resources GILS records describe in accordance with the categories shown in Table E2-2.

Table E2-2
Aggregation Semantics

Code	Operational Definition	Examples
Record Aggregates Objects	GILS record, by virtue of its creation, collects discrete information resources that record content indicates would not have otherwise been collected or aggregated. Assigned in the absence of clues within the record that the represented objects were heretofore packaged as <i>this collection</i> to optimize information discovery and retrieval.	<ul style="list-style-type: none"> • Privacy Act Systems compilation • files • press releases • forms
Aggregated Object Represented	GILS record represents an <i>a priori</i> or purposeful collection of information resources—e.g., woodpecker database or agency website. GILS record represents an object that collects, or comprises, two or more discrete information objects, and that represents a collection of standalone information files or products packaged together on the basis of a common theme or subject for functional convenience.	<ul style="list-style-type: none"> • CD-ROM of regulations • System that compiles Privacy Act records • job line of open requisitions
Discrete Object Represented	GILS record describes a standalone document-level entity that does not meet the criteria for “object aggregates metadata” below.	<ul style="list-style-type: none"> • annual report • videotape
Object Aggregates Metadata	GILS record describes a pre-existing metadata collection, or “locator,” as an information resource.	<ul style="list-style-type: none"> • directory • catalog • index • log

As shown in the following data summaries, “subject matter databases” (i.e., a database of endangered species) were often described, comprising one-quarter of the records sampled. The Core subset sample reflected a higher value for “publication”—a category including discrete information objects available as self-contained entities such as books and individual technical reports, etc.

TOTAL SAMPLE

OBJECT REPRESENTED	N	%
SUBJECT MATTER DATABASE	18	22%
PUBLICATION	16	19%
MISCELLANEOUS DOCUMENTS IN <i>AD HOC</i> COLLECTION	14	17%
AGENCY HOMEPAGE	8	10%
ORGANIZATION	6	7%
FORM	4	5%
ADMINISTRATIVE CATALOG	3	4%
BIBLIOGRAPHIC DATABASE	3	4%
PUBLICATIONS CATALOG	4	5%
SYSTEM OF SYSTEMS	3	4%
PROGRAM	2	2%
JOB LINE	1	1%
UNKNOWN	1	1%
TOTAL	83	100%

CORE SUBSET

OBJECT REPRESENTED	N	%
PUBLICATION	11	26%
SUBJECT MATTER DATABASE	8	19%
MISCELLANEOUS DOCUMENTS IN <i>AD HOC</i> COLLECTION	7	17%
AGENCY HOMEPAGE	4	10%
ORGANIZATION	2	5%
FORM	2	5%
PUBLICATIONS CATALOG	2	5%
SYSTEM OF SYSTEMS	2	5%
ADMINISTRATIVE CATALOG	1	2%
BIBLIOGRAPHIC DATABASE	1	2%
PROGRAM	1	2%
JOB LINE	1	2%
UNKNOWN	0	0%
TOTAL	42	100%

Discussion and Recommendations: The findings above show that nearly 20% of records sampled described a discrete publication as opposed to a "locator" (e.g., catalog) resource. This perhaps indicates an incorrect assumption by policymakers that all of an agency's "information dissemination products" (including any book, paper, map, machine-readable material, audiovisual production, or other documentary material, regardless of physical form or characteristic, disseminated by an agency to the public") were described by extant locators. In fact, "locator-function" objects (extrapolated from Administrative Catalog, Bibliographic Database, and Publications Catalog) represented only a little more than 10% of objects described in the sample. See C3.1-Title, C3.4-Abstract, C3.6-Agency Program, C4.4-Availability-Resource Description, and P6-Aggregation for implications and recommendations associated with identifying "objects" within GILS records. In summary, it is highly recommended that users be provided an additional GILS element of "object represented" in order to evaluate aggregation, or "informational distance" from satisfaction of their requirement, and to increase retrieval precision (e.g., by eliminating object type "job line" from a search on unemployment statistics).

P3 Record Aggregation

Findings: The analysis attempted to characterize each sampled record's "aggregation" or "granularity" relative to others in the sample. Appendix C-4 Record Content Analysis Methodology provides a complete discussion of semantics; however, a summary of definitions is provided below.

- *Record aggregates object.* The GILS record, by virtue of its creation, collects discrete information resources that the record content indicates would not have otherwise been collected or aggregated—e.g., "Privacy Act Systems," "General Files", "Press Releases", or "Forms". This code was assigned in the absence of clues within the record that the represented objects were heretofore packaged *as this collection* to optimize information discovery and retrieval.
- *Aggregated object represented.* The GILS record represents an *a priori* or purposeful collection of information resources—e.g., "Woodpecker Database" or an agency Web site. In other words, the GILS record represents an object that collects, or comprises, two or more discrete information objects, and that object represents a collection of standalone information files or products packaged together on the basis of a common theme or subject for functional convenience—e.g., a CD-ROM of regulations, a discrete system of Privacy Act records, and a voice recording of employment opportunities.
- *Discrete object represented.* The GILS record describes a standalone document-level entity that does not meet the criteria for "object aggregates metadata" below—e.g., an Annual Report or videotape.
- *Object aggregates metadata.* The GILS record describes a pre-existing metadata collection, or "locator"—e.g., directory, catalog, index, or log—as an information resource.

The analysis revealed a high number of records, more than one-third of the sample, that appeared to aggregate records, document, files, and other objects. Attachment 2a, a record titled "Briefing Materials, Public Comments, Other Related Official Files" exemplifies this phenomenon.

Records describing a discrete information resource that comprises two or more sub-resources (aggregated object represented) constituted 25% and 31% of the total and Core subset sample, respectively. Attachment 2b, a record titled "Worldwide Real Property Inventory System," describes a system that aggregates data from other systems.

Approximately one-fourth of the records sampled described a discrete, standalone information object such as a book, video, or technical report (see Attachment 2c "Investment Fund Brochure"). Metadata collections were described about 1 in 10 records; Attachment 2d, a record titled "Office of the General Counsel Library Catalog," serves as an example of items in this category. Five records whose "objects represented" (see P5) were not information resources (e.g., "organizations," "[functional] programs," and "facilities") could not be coded as to aggregation.

TOTAL SAMPLE

AGGREGATION	N	%
RECORD AGGREGATES OBJECTS	30	36%
AGGREGATED OBJECT REPRESENTED	21	25%
DISCRETE OBJECT REPRESENTED	17	20%
OBJECT AGGREGATES METADATA	10	12%
UNKNOWN	5	6%
TOTAL	83	100%

CORE SUBSET

AGGREGATION	N	%
RECORD AGGREGATES OBJECTS	16	38%
AGGREGATED OBJECT REPRESENTED	13	31%
DISCRETE OBJECT REPRESENTED	7	17%
OBJECT AGGREGATES METADATA	4	10%
UNKNOWN	2	5%
TOTAL	42	100%

Discussion and Recommendations: More than one-third of records described "new collections"—i.e., resources aggregated by virtue of the record itself, although the serviceability of this phenomenon in NIDR is questionable given record titles such as "Minutes" and "General Files." Counterbalancing this finding, however, is a roughly 20% incidence of item-level description when both "record aggregates metadata" (e.g., a catalog is the object) and "discrete object represented" (e.g., a brochure is the object). Investigators found the task of characterizing granulation extremely challenging and recognize the problems of nonexclusivity and intercoder reliability in the method employed during this analysis. However, a similar coding scheme, perhaps based on clearly defined (including by way of example) steps of distance from satisfaction of an information need and deliberately associated with object type, may be useful in facilitating NIDR. Section C4.4-Availability-Resource Description addresses the issue of granularity as well.

P4 Containers

Findings: "Containers" were defined as dissemination media. Nearly half of the total sample described information resources whose containers could not be discerned from record content; containers in the Core subset were more frequently mentioned (31%). More than 1 in 4 of both the total and subset records described resources available in multiple containers. "Print" resources comprised about 20% of both samples, followed by "Web" resources (about 10%).

TOTAL SAMPLE

CONTAINER	N	%
UNKNOWN	34	41%
MULTIPLE (OF ALL)	18	22%
PRINT	19	23%
WEB	7	8%
DIALUP	2	2%
CD-ROM	1	1%
VOICE	1	1%
VIDEO	1	1%
RADIO/TV	0	0%
BROADCAST		
MICROFORM	0	0%
LISTSERVE	0	0%
GOPHER	0	0%
FTP	0	0%
FAX	0	0%
EMAIL	0	0%
TOTAL	83	100%

CORE SUBSET

CONTAINER	N	%
UNKNOWN	13	31%
MULTIPLE (OF ALL)	12	29%
PRINT	10	24%
WEB	4	10%
DIALUP	1	2%
CD-ROM	1	2%
VOICE	1	2%
VIDEO	0	0%
RADIO/TV	0	0%
BROADCAST		
MICROFORM	0	0%
LISTSERVE	0	0%
GOPHER	0	0%
FTP	0	0%
FAX	0	0%
EMAIL	0	0%
TOTAL	42	100%

Discussion: User recognition of and ability to access/use "containers" may be a significant factor in relevance judgment (see discussion at C4.4-Resource Description). While time constraints precluded an assessment of hypertext incidence within Web containers, S2-Content Hypertext provides data concerning total incidence within the sample.

Recommendations: See C4.4-Resource Description.

4.4. Serviceability

The "serviceability" data shown below are considered to represent record effectiveness in terms of the degree to which they enhance NIDR, convenience to the user, aesthetics, readability, and relevance judgment.

S1 File Formats

Findings: The HTML file format was available for nearly half of the total records sampled and 60% of records in the Core subset. Two records were available in SGML; none were available in PDF.

TOTAL SAMPLE

ASCII TEXT AVAIL?	N	%	HTML AVAIL?	N	%	PDF AVAIL?	N	%	SGML AVAIL?	N	%
YES	67	81%	YES	38	46%	YES	0	0%	YES	2	2%
NO	16	19%	NO	45	54%	NO	83	100%	NO	81	98%
TOTAL	83	100%	TOTAL	83	100%	TOTAL	83	100%	TOTAL	83	100%

CORE SUBSET

ASCII TEXT AVAIL?	N	%	HTML AVAIL?	N	%	PDF AVAIL?	N	%	SGML AVAIL?	N	%
YES	32	76%	YES	25	60%	YES	0	0%	YES	2	5%
NO	10	24%	NO	17	40%	NO	42	100%	NO	40	95%
TOTAL	42	100%	TOTAL	42	100%	TOTAL	42	100%	TOTAL	42	100%

Discussion: Each file format has advantages for record creators and users, however given a choice between HTML- and ASCII-formatted records, study participants expressed a clear preference for the former (see Appendix E-3 Scripted Online User Assessment Findings, Discussion, and Recommendations). See also A1-Format Errors. Also, it is noted that the use of the word "TEXT" vs. "HTML" on results lists may lead users to believe that clicking on the former will provide the (full) text of a document.

Recommendations: As noted in other sections, investigators believe that standardization of record display will contribute to users' "footing" within GILS vs. other information—e.g., recognition that GILS is a bounded (by function), top-down, two-dimensional service that spans across all agencies. The ready availability of "HTML editor" programs, which convert a variety of file formats to HTML, should be exploited as resources permit.

S2 Content Hypertext

Findings: Roughly 1 in 4 of the records sampled featured at least one instance of hypertext somewhere within element values. The following tables show that Available Linkage and Distributor Network address were the most frequently hotlinked elements (N = number of records). The incidence of Cross-Reference element hypertext was negligible.

TOTAL SAMPLE

ELEMENTS WITH HOTLINKED VALUE	N	% OF TOTAL USE
AVAILABLE LINKAGE	13	52%
DISTRIBUTOR NETWORK ADDRESS	6	24%
DISTRIBUTOR URL (LOCALLY DEFINED)	2	8%
ABSTRACT	1	4%
CROSS-REFERENCE LINKAGE	1	4%
CROSS-REFERENCE TITLE	1	4%
ORDER PROCESS	1	4%
TOTAL	25	100%
NUMBER OF RECORDS USING LINKS	19	23%

CORE SUBSET

ELEMENTS WITH HOTLINKED VALUE	N	% OF TOTAL USE
AVAILABLE LINKAGE	13	52%
DISTRIBUTOR NETWORK ADDRESS	6	24%
DISTRIBUTOR URL (LOCALLY DEFINED)	2	8%
ABSTRACT	1	4%
CROSS-REFERENCE LINKAGE	1	4%
CROSS-REFERENCE TITLE	1	4%
ORDER PROCESS	1	4%
TOTAL	25	100%
NUMBER OF RECORDS USING LINKS	12	29%

Discussion: While the maintenance burden of hypertext is recognized, users' expectation for it will continue to accelerate for the foreseeable future.

Recommendations: A further analysis of hypertext incidence and placement is warranted to capitalize on user expectations. In the Scripted Online User, for example, users were asked "What do you think would happen if you were to click on this record's hypertext title? ["U.S. International Trade Commission. Library Services"], and presented a multiple-choice list of:

- a. I would jump to the ITC website
- b. I would connect to ITC's online library catalog
- c. A list of ITC library staff contacts would appear
- d. I would link to a fuller/longer version of this record
- e. I would be given a list of library services such as interlibrary loan, photocopying, and research assistance
- f. Other (please specify)

It is recommended that system designers actively participate in and contribute to PURL and similar technology development efforts to ensure satisfaction of GILS unique requirements.

S3 Capitalization

Findings: 86% of records sampled used sentence-case capitalization; 10% used all capitals for element labels.

TOTAL SAMPLE

CAP STYLE	N	%
SENTENCE CASE	71	86%
ELEMENTS ONLY CAP	8	10%
MIXED	4	5%
OTHER	0	0%
ALL CAPS	0	0%
TOTAL	83	100%

CORE SUBSET

CAP STYLE	N	%
SENTENCE CASE	35	83%
ELEMENTS ONLY CAP	5	12%
MIXED	2	5%
OTHER	0	0%
ALL CAPS	0	0%
TOTAL	42	100%

Discussion and Recommendations: Attachments to this appendix, prepared to support Sections 1.5 Examples of High Quality Records and the discussion of P3-Aggregation, represent some variations in GILS record appearance. Investigators believe that standardization of record display, including typeface and weight will contribute to users' "footing" within GILS vs. other information space—e.g., recognition that GILS is a bounded (by function), top-down, two-dimensional service that spans across all agencies. It is recommended that decisions concerning standardization be based on published research concerning visual cues in human-machine interaction.

S4 Indentation

Findings: Roughly 1 in 4 records featured all flush-left text —i.e., no indentation was used to represent the element/subelement hierarchy.

TOTAL SAMPLE

INDENTATION	N	%
YES	61	73%
NO	22	27%
TOTAL	83	100%

CORE SUBSET

INDENTATION	N	%
YES	29	69%
NO	13	31%
TOTAL	42	100%

Discussion and Recommendations: The lack of indentation, when coupled with other style characteristics, such as no boldface, all capitals, etc., impedes scanning of record content for relevant terms. Investigators believe that standardization of record display, including indentation, will contribute to users' "footing" within GILS vs. other information space—e.g., recognition that GILS is a bounded (by function), top-down, two-dimensional service that spans across all agencies. It is recommended that decisions concerning standardization be based on published research concerning visual cues in human-machine interaction.

S5 Element Display Order

Findings: Roughly two-thirds of the total sample and nearly one-half of the Core subset records displayed elements in the order recommended by *FIPS Pub. 192* and the NARA *Guidelines*.

TOTAL SAMPLE

PREFERRED DISPLAY ORDER	N	%
YES	53	64%
NO	30	36%
TOTAL	83	100%

CORE SUBSET

PREFERRED DISPLAY ORDER	N	%
YES	24	57%
NO	18	43%
TOTAL	42	100%

Discussion and Recommendations: Lack of consistent and predictable ordering of metadata elements is disconcerting to users and may inhibit recognition of relevant terms. It is recommended that record designers standardize and use an ordering scheme based on systematic analysis of various user communities' preferences and consider optional (on-the-fly) re-ordering or suppression of elements upon client command as is provided by Z39.50-compliant servers. See also C1-Number of Elements Per Record.

S6 Definition of Acronyms

Findings: This analysis considered incidence of acronyms anywhere in the record, including the Acronym subelement. Only 12% of records containing acronyms failed to define them.

TOTAL SAMPLE

ACRONYMS	N	%
DEFINED	19	23%
NOT DEFINED	10	12%
NOT USED	54	65%
TOTAL	83	100%

CORE SUBSET

ACRONYMS	N	%
DEFINED	13	31%
NOT DEFINED	5	12%
NOT USED	24	57%
TOTAL	42	100%

Discussion and Recommendation: The incorporation of defined acronyms in government information undoubtedly assists users in NIDR. The absence of acronyms altogether in more than half of the records sampled was a surprising result, and record creators should be trained not to sacrifice relevant acronyms for record brevity. Section C3.1-Titles addresses the use of AIS acronyms.

S7 Citation of Legislation

Findings: Roughly half of the records sampled contained a reference to legislation concerning the information resource and/or its provision, including one instance of reference to GILS in Agency Program.

TOTAL SAMPLE

LEGISLATIVE CITE	N	%
YES	40	48%
NO	43	52%
TOTAL	83	100%

CORE SUBSET

LEGISLATIVE CITE	N	%
YES	21	50%
NO	21	50%
TOTAL	42	100%

Discussion and Recommendation: Inclusion of references to pertinent legislation improves NIDR only when the user searches in the "correct" (as used by the record creator) format—e.g., "Pub. L. 103-40" vs. "public law 103"). Given the fairly high incidence of legislative citation, it is recommended that a qualitative analysis of the citation formats be performed to determine the extent of variations. If warranted, a standard format should be incorporated into the NARA *Guidelines*. The feasibility of automating the insertion of substantively correct and properly formatted legislative citations upon inclusion of a program or project name (i.e., installation of an expert-system legislation index cross-reference macro) could reduce record creation and maintenance burden as well as facilitate identification of legislation-dependent resources for agency users (IRM and public information office personnel), Congressional users, and public policy researchers.

S8 Locally Defined Elements

Findings: This analysis considered locally defined elements as any not specifically mentioned in the NARA *Guidelines*. The 12 identified (labeled) locally-defined elements are listed below. (Note: some values were null.)

- URL
- System Products Disposition
- Organization [appears between Title and Abstract in addition to and of equal weight with Originator]
- Material Type
- Creation Date
- Disposition
- [Agency] Storage Authorized
- Disposition Authority
- Description [sic]
- Record-Type
- Status
- Date Of Last Review

Discussion and Recommendations: The NARA *Guidelines* provide for use of locally defined elements in organizational elements and "when agencies wish to convey to the public or use for internal purposes information that is not part of the GILS Core. Whenever possible, such supplemental information should be associated with one of the GILS Core Elements as a locally-defined subelement to the Core Element [rather than in Supplemental Information]." It is recommended that the incidence and values of locally-defined elements be studied systematically; high incidence may indicate the need for an additional "official" element and extensive usage of a locally defined element that definitionally belongs with an extant element may dictate a clarification of the NARA *Guidelines* usage instructions.

5.0. SUMMARY OF RECOMMENDATIONS

The following recommendations have been extracted from Section 4.0 Findings, Discussion, and Recommendations and arranged according to whether they are best implemented (1) at the agency level or (2) standards level by means of revision to the NARA *Guidelines*.

5.1. Agency Level

1. Devise a hard-/software independent template and/or HTML editor for record formatting, or limit formatting responsibility to agency or subcontracted personnel with Web browsers.
2. Use machine-based spell checkers, or assign checking responsibility to someone other than the writer.
3. Use "auto-correct/auto-format" macros to standardize dates.
4. Audit the accuracy of Distributor and Points of Contact element values.
5. Enlist an objective, third-party (i.e., other than resource creator and/or record creator) to evaluate Title descriptiveness.
6. Automate the insertion of substantively correct and properly formatted legislative citations upon inclusion of a program or project name (i.e., install an expert-system legislation index cross-reference macro).
7. Develop training and complementary procedures for record creators to recognize the need for and to obtain the information for populating the Methodology, Sources of Data, and Schedule Number elements.
8. Use computer-generated time/stamp dates as a final "record checked and approved for release" procedure to promote responsibility for record content quality, facilitate audit sampling, and ease record maintenance.
9. Provide record creators and quality checkers a short, straightforward, procedural set of record-creation instructions as well as a FAQ list, pocket guide reference, context-sensitive online help, etc. to select elements required to describe the resource and its availability appropriately.
10. Monitor the incidence and values of new or "truly" locally-defined as well as those more correctly used as an extant element subelement to inform revision of the NARA *Guidelines*.
11. Implement Z39.50-compliant servers and clients, which will present customized views of the record through re-ordering or suppression of elements upon client command (e.g., allowing the user to select presentation of the Abstract with the Title in the results list to assist in relevance judgment).
12. On GPO Access GILS, provide a hypertext link from "US Government Manual" in the field definition files to a recast version of *Government Manual* Appendix A: Commonly Used Abbreviations and Acronyms
13. Participate in and pursue PURL and similar technology development efforts to ensure satisfaction of GILS unique requirements

5.2. Standards Level (Revision of the NARA *Guidelines*)

1. Content analysis of the *Guidelines* by a third-party (i.e., not someone who was involved with the writing) is recommended to reveal areas of redundancy and ambiguity—e.g.,
 - a. Restate the *Guidelines* Distributor element requirement from "Complete as many of the subordinate fields as necessary to identify the party" to "Complete all subordinate fields as available to assist the user in communicating with the Distributor."
 - b. More clearly differentiate between placement of certain constraints (e.g., "requires forms-capable browser" in Access Constraint vs. Order Process)
 - c. More clearly differentiate expected values between Purpose and Agency Program
 - d. Clarify the rationale for isolating mandatory use of Sources of Data to describe AISs
2. Rename the Date of Last Modification element "Record Revision Date" and eliminate confusing "entry" nomenclature in the usage guidelines
3. Revise the definition, usage, and structure of the Resource Description subelement to provide physical description of the object as recognizable by the *user* rather than by the distributor; move this information

from the Abstract element to the Resource Description subelement; and make the (redefined) Resource Description subelement mandatory and structurally associated with the Abstract element.

4. Provide the additional GILS element "Object Represented" in order to evaluate aggregation, or "informational distance" from satisfaction of the information requirement and increase retrieval precision.
5. Standardize record display, including type font, weight, and size, as well as indentation and capitalization to "moor" users in GILS information space and promote the concept of a government-wide rather than agency-centric program. Base decisions on recognized research concerning visual cues in human-machine interaction.
6. Standardize and use an element display order scheme based on systematic analysis of various user communities' preferences.
7. Re-examine the concept and functionality of the GILS "Core." Should the requirement remain viable, clarify the rationale and guidance.

It is recommended that agencies and inter-agency oversight bodies implement the above recommendations with specific procedures, schedules, and performance measurements.

6.0. FURTHER RESEARCH

Investigators discovered from this content analysis experience that the method might find optimal utility when employed in circumstances where specific user-defined criteria are known. For example, the user assessment technique described in Appendix C-5 Scripted Online User Assessment Methodology, for example, revealed that excessive record length (i.e., number of elements present) was off-putting to some users. The effects of record length on users for known-item searching vs. browsing, for example, could be studied by means of a record sample comprising a mix of popular resources (determined by Public Information Officers or transaction log analysis), newly created or newly aggregated resources, and resources sought frequently by professional intermediaries (such as federal depository librarians) but not end users. An understanding of how users read, evaluate, and "use" GILS records could inform the creation of customized record views.

The following list presents areas for further research of record content as an indicator of how well GILS is meeting expectations of users.

1. Hypertext incidence and placement (see Appendix C-5 and Appendix D-5 Online Scripted User Assessment Methodology and Instrument, respectively)
2. Government-wide frequency of change and clustering of Dates of Last Modification
3. Effects of "familiar" (e.g., LCSH) and specialized Controlled Vocabularies on NIDR *vis-a-vis* users' dependency on knowledge of agency mission and to increase precision of information retrieval (this effort should be in concert with the Library of Congress and GPO's Cataloging Branch)
4. User preferences for "catalog- or browsing based" NIDR (e.g., the approach of Yahoo! and the Argus Clearinghouse)
5. Appropriateness of content and placement of element values; a possible research project might isolate a random sample of selected data values and test whether GILS-cognizant vs. noncognizant users can place them in the "correct" (per published standards) elements
6. Record and resource aggregation effects on NIDR in terms of distance from satisfaction of an information need.
7. Control Identifier values as representing resource accession, IRM, or subject/object classification schemes.
8. Content of Web pages linking to GILS' homepages as a means to improve Local Subject Index Terms and Cross References through increased understanding of user expectations concerning the scope or nature of the GILS record collection. For example, links predominately via Web sites of the legal community may permit inclusion of more specific legal terms as well as citation of applicable cross-references. One current method of acquiring linking-from data is to execute an AltaVista [search along the lines of link:](http://www.altavista.digital.com) [http://www.\[agency name\].gov/gils](http://www.[agency name].gov/gils) - host:[http://www.\[agency name\].gov](http://www.[agency name].gov). (Instructions are provided at the Alta Vista site.).

9. Evolution of data elements over time (as documented in working group minutes, listservs, standards, user-based research results, and agency-level training materials and procedures) as elucidating drivers for change—e.g., international, Federal, and state information policy; technology; standards; economics; nature of the resources; information life cycles; user expectations; agency mission; etc.
10. Content analysis of maintenance-intensive metadata (e.g., organizational subelements, Availability element URLs, cost, etc.) to reveal means of consolidating or otherwise arranging such data for ease or possible automation of update or record archiving.
11. Incidence of (1) a single information resource being “claimed” by multiple Record Sources—e.g., a CD-ROM content creator, content compiler, manufacturer, distributor and (2) “nonoriginal” Control Identifiers as an indicator of need for consensual policies on record-creation responsibility to avoid duplication of efforts and unnecessary maintenance burden.

These areas of research are considered viable in terms of improving GILS synergy with other products and services in government information space and its efficacy in NIDR, as well as in developing technical and procedural standards to guide the evolution of GILS.

7.0. CONCLUSION

The content analysis of GILS records provided valuable data concerning their accuracy, completeness, and serviceability in NIDR. In addition, it afforded a high-level view of the types and aggregation levels of resources being described as well as modes of packaging and dissemination.

Placed in context with the online user assessment results (see Appendix E-3 Scripted Online User Assessment Findings, Discussion, and Recommendations), the assessed degree of variation in the quality of GILS records demands immediate attention at both the agency and standards levels. Agencies can adapt the content analysis methodology to appraise the quality of their current GILS records as well as to identify early and act on evolving issues—e.g., such as scope of collection and levels of granularity—at the collection level. In addition, when applied to a government-wide sample, adaptations of this record content analysis will inform oversight bodies about levels of adherence to standards and synergy of GILS with other Federal government NIDR mechanisms.

Attachments E2-1a through E2-1d Examples of High-Quality Records From the Sample

Four records exhibiting characteristics of "high quality" as defined by the record content analysis are reproduced in Attachments 1a through 1d. They have been reproduced as closely as possible to depict their actual display to an online user of GILS (i.e., font attributes, line spacing and indentation, etc.):

- *AHCPR Publications Clearinghouse* available at <<http://www.dhhs.gov/progorg/oirm/newhhsgils.htm>> by searching Control Identifier (quotes required) "HHS-AHC-00509"
- *Aviation Accident Synopses World Wide Web Page* available by browsing <<http://www.ntsb.gov/Info/Info.htm>> or directly at <<http://www.ntsb.gov/Info/GILS/GILSSYN.htm>>
- *Farm Credit Administration's Privacy Act Systems* available <http://www.access.gpo.gov/su_docs/gils/gilsfld.html> by searching Control Identifier (quotes required) "FCA/PA-1"
- *FEMA Publications Catalog* available <http://www.access.gpo.gov/su_docs/gils/gils.html> by searching Federal Emergency Management Agency for "FEMA0001"

These records contain mandatory elements populated with NARA *Guidelines*-compliant values and are highly readable and descriptive without excessive length. In addition, they represent a range of "information objects" and "containers": an information resource organization, an aggregated set of reports available via Web site, a "system of records" available via GPO Access, and a traditional printed publications catalog, respectively.

**Attachment E2-1a
High Quality Record From the Sample**

Title: AHCPR Publications Clearinghouse

Acronym: AHCPRPC

Originator: Department of Health and Human Services (HHS)
U.S. Public Health Service (PHS)
Agency for Health Care Policy and Research (AHCPR)

Local Subject Index: US Federal GILS, Clearinghouse

Abstract: The AHCPR Publications Clearinghouse is the primary storage and distribution point for all AHCPR publications. The Clearinghouse also maintains and manages AHCPR's automated mailing/inventory control system and manages the storage and shipping of AHCPR exhibits.

Purpose: The AHCPR Publications Clearinghouse serves as the Agency's direct mail and fulfillment service point.

Agency Program: A component of the Center for Health Information Dissemination, this clearinghouse is authorized under Section 6013, P.L. 101-239, Omnibus Reconciliation Act of 1989.

Time Period of Content:

Time Period Structured: 19940901 -

Time Period Textual: 1 September 1994 - ongoing

Availability:

Distributor:

Name: Erin Henderson, Project Director

Organization: Logistics Applications, Inc.

Street Address: 9475 Gerwig Lane, Suite V

City: Columbia

State: MD

Zip Code: 21046-1506

Country: USA

Network Address: None

Hours of Service: 9 a.m. - 5 p.m., Monday - Friday
closed Federal Holidays

Telephone: 800-358-9295; 301-621-3033 (local and non-U.S.)

Fax: 410-290-3841

TDD: 888-586-6340

Order Process: AHCPRPC can be accessed via telephone, fax, TDD, or the Internet via the AHCPR Web Home Page.

Technical Prerequisites: For WWW access, Internet access and a browser is required.

Available Time Period:

Electronic information access: 1 September 1994 - ongoing

Available Linkage:

Connect to AHCPRPC via AHCPR Home Page:
<http://www.ahcpr.gov/>

Available Linkage Type: plain text

Access Constraints: None

Use Constraints: None

Point of Contact:

Name: Steven Merrill, Federal Project Officer;
Judy Wilcox, Federal Alternate Project Officer
Organization: Agency for Health Care Policy and Research
Street Address: 2101 E. Jefferson Street, Suite 501

City: Rockville

State: MD

Zip Code: 20852

Country: USA

Network Address: smerrill@po5.ahcpr.gov;
awilcox@po5.ahcpr.gov or
puborder@po5.ahcpr.gov

Hours of Service: Monday - Friday, 7 a.m. - 4 p.m.

Telephone: Steven Merrill: 301-594-1364, extension 1350;
Judy Wilcox: 301-594-1364, extension 1389

Fax: 301-594-2286

Control Identifier: HHS-AHC-00509

Record Source: Agency for Health Care Policy and Research
Center for Health Information Dissemination

Date of Last Modification: 19960708

**Attachment E2-1b
High Quality Record From the Sample**

Title: Aviation Accident Synopses World Wide Web Page

Originator:

Department/Agency Name: National Transportation Safety Board

Major Organizational Subdivision: Office of Research and Engineering

Name of Unit: Analysis and Data Division

Local Subject Index:

Local Subject Term: US Federal GILS

Abstract: The Aviation Accident Synopses World Wide Web Page contains short reports describing aircraft accidents and incidents and their probable cause, and contributing factors. Included are civil aviation accidents within the United States, its territories and possessions, and in international waters. Incidents (mishaps that do not meet the aircraft damage or personal injury thresholds in the regulatory definition of "accident") investigated by the National Transportation Safety Board are contained in the database in the same form as accidents. Reports are available for the time period 1983 to the present. Generally, a preliminary report is available on line within a few days of an accident. When the investigation is completed, the preliminary report is replaced with a final description of the accident and its probable cause. The World Wide Web page provides access to more than 35,000 reports through 1995, and is growing by approximately 2,250 cases per year. Access to specific accidents and incidents is by means of monthly lists of all such occurrences in the National Transportation Safety Board's Aviation Accident Data Base. Synopses and monthly lists are updated daily.

Purpose: The Aviation Accident Synopses World Wide Web page provides to the public direct access to limited information regarding aviation accidents investigated by the National Transportation Safety Board.

Agency Program: The National Transportation Safety Board provides information to the public via its World Wide Web page with the intent of making safety related information rapidly and easily available to its customers worldwide.

Spatial Reference:

Geographic Name:

Geographic Keyword Name: United States

Geographic Keyword Name: United States Territories and Possessions

Geographic Keyword Name: International Waters

Time Period of Content:

Time Period-Structured: 1983-

Time Period-Textual: 1983 - [ongoing]

Availability:

Distributor:

Name: NTSB WWW Server

Organization: National Transportation Safety Board

Network Address: webmaster@ntsb.gov

Order Process: Accident synopses are available via the internet using an http client. From the NTSB Home Page (www.ntsb.gov), select "Aviation" then "Accidents".

Technical Prerequisites: Access to the Internet and an http client.

Available Linkage: <http://www.ntsb.gov>

Available Linkage Type: text/plain

Sources of Data: Synopses are produced from data developed by aviation accident investigators of the National Transportation Safety Board.

Access Constraints: Synopses may be located only by searching monthly lists of accidents and incidents.

Use Constraints: Public Law 93-633 and the National Transportation Safety Board's regulations 49 CFR 835 prohibit the use of accident/incident findings, including the probable cause and contributing factors as evidence in any suit or action for damages arising from that event.

Point of Contact:

Name: Analysis and Data Division (RE-50)

Organization: National Transportation Safety Board

319

Street Address: 490 L'Enfant Plaza East

City: Washington

State: DC

Zip Code: 20594-2000

Country: USA

Network Address: webmaster@ntsb.gov

Hours of Service: 8:00 a.m. - 5:00 p.m.

Telephone: 202-314-6550

Telephone: 800-877-6799

Fax: 202-314-6598

Cross Reference:

Cross Reference Title: Aviation Accident Database

Cross Reference Linkage: GILS: NTSB0001

Cross Reference Type: text/html

Cross Reference Title: Safety Recommendations Database

Cross Reference Linkage: GILS: NTSB0002

Cross Reference Type: text/html

Cross Reference Title: Accident Investigation Dockets

Cross Reference Linkage: GILS: NTSB0003

Cross Reference Type: text/html

Cross Reference Title: Formal Reports

Cross Reference Linkage: GILS: NTSB0004

Cross Reference Type: text/html

Schedule Number: Not scheduled

Control Identifier: NTSB0005

Record Source:

Department/Agency Name: National Transportation Safety Board

Major Organizational Subdivision: Office of Research and Engineering

Name of Unit: Analysis and Data Division

Date of Last Modification: 19960329

**Attachment E2-1c
High Quality Record From the Sample**

TITLE: Farm Credit Administration's Privacy Act Systems

Originator: Farm Credit Administration

LOCAL SUBJECT INDEX: US Federal GILS

ABSTRACT: In compliance with the Privacy Act of 1974, the Farm Credit Administration publishes notices in the Federal Register about the record systems the agency maintains that are retrieved by name or personal identifier. These record systems are commonly referred to as "Privacy Act systems" and the information published about them are referred to as "system notices". The Privacy Act systems maintained by FCA primarily cover FCA employees. Each "system notice" contains the following information: system name; system location; categories of individuals covered by the system; categories of records in the system; authority for maintenance for the system; routine use of records maintained in the system; policies and practices for storing, retrieving, accessing, retaining, and disposing of records in the system; the name and address of the system manager; and procedures for access and contesting information in the records. After publication in the Federal Register, FCA's system notices are compiled in Privacy Act Issuances, a biennial compilation of all Federal agency Privacy Act notices, which has been published by the Government Printing Office in CD-ROM format since 1993. A free public-access version of Privacy Act Issuances is available on the Internet. Any revisions to FCA's Privacy Act notices since the last compilation can be found by reviewing the Federal Register.

PURPOSE: FCA's Privacy Act system notices provide the public with information about systems of records maintained by the agency that are retrieved by name or personal identifier.

AGENCY PROGRAM: 5 USC 552a, the Privacy Act of 1974

AVAILABILITY:

Distributor:

Name: Superintendent of Documents, U.S. Government Printing Office (GPO)

Street address: Mail Stop SDE, 732 North Capitol Street NW

City: Washington, D.C

Zip Code: 20401

Country: USA

Telephone: 202-512-1530

Fax: 202-512-1262

Hours: 7:00 a.m. - 5:00 p.m. M-F EST

Resource: Privacy Act Issuances (on-line version)

Order Process: Available through the Internet free of charge or through dial-up access. There is no charge for using the phone line, but you may incur long distance charges.

Technical Prerequisites: Access to Internet and an http browser or telnet or WAIS client or access to a modem and telephone line.

Linkage: <http://www.access.gpo.gov/sudocs/gils/gils.html>; telnet: swais.access.gpo.gov; wais: wais.access.gpo.gov; asynchronous: 202- 512-1661

Linkage Type: text/plain

ACCESS CONSTRAINTS: None.

USE CONSTRAINTS: None.

323

POINT OF CONTACT:

Name: Office of General Counsel

Organization: Farm Credit Administration

Street address: 1501 Farm Credit Drive

City: McLean

State: Virginia

Zip code: 22102-5090

Country: USA

Telephone: 703-883-4022

CONTROL IDENTIFIER: FCA/PA-1

RECORD SOURCE: Farm Credit Administration

DATE OF LAST MODIFICATION: 19970220

**Attachment E2-1d
High Quality Record From the Sample**

Title: FEMA Publications Catalog

Originator:

Department/Agency Name: Federal Emergency Management Agency

Abstract: This catalog provides information regarding publications and other printed matter produced by FEMA.

Purpose: This catalog provides directions/addresses of where and how to obtain copies of FEMA publications and other printed matter produced by FEMA.

Agency Program: These publications are made available as a customer service to the public. The publications also document FEMA's mission and programs that are available to the public, such as emergency management training.

Availability:

Distributor:

Distributor Name: Federal Emergency Management Agency
Distributor Organization: FEMA Publications Warehouse
Distributor Street Address: P. O. Box 2012
Distributor City: Jessup
Distributor State: MD
Distributor Zip Code: 20794-2012
Distributor County: USA
Distributor Network Address: None
Distributor Hours of Service: 8:30 a.m. to 5:00 p.m. EST
Distributor Telephone: 1-800-480-2520
Distributor Fax: 301-497-6378

Order Process: Requests for copies of this catalog can be made by writing or calling and asking for FEMA Publication 20. Requests are limited to 100 copies. Any of the publications listed in the catalog can be ordered by following the instructions listed in the catalog.

Access Constraints: There are no access constraints for the catalog for requests from the United States; however, individual publications listed in the catalog may indicate some access constraints. Foreign requests must be approved by the Office of Security, FEMA, prior to being fulfilled.

Use Constraints: None.

Point of Contact:

Contact Name: Printing & Publications Branch
Contact Organization: Program Services Division, FEMA
Contact Street Address: 500 C Street, SW
Contact City: Washington
Contact State: DC
Contact Zip Code: 20472
Contact Country: USA
Contact Network Address: None
Contact Hours of Service: 8:30 a.m. to 5:00 p.m. EST
Contact Telephone: 202-646-2650
Contact Fax: 202-646-3524

Schedule Number: Scheduled: N1-311-86-1 1K6

Control Identifier: FEMA0001

Record Source: Federal Emergency Management Agency

Date of Last Modification: 19950404

326

Attachments E2-2a through E2-2d Aggregation Code Examples

These attachments illustrate the operational definitions of record and resource aggregation used in the record content analysis; no other evaluative criteria apply.

The following four aggregation codes, discussed more fully in Appendix C-4 Record Content Analysis Methodology, were applied to all sampled records:

Code	Operational Definition	Examples
Record Aggregates Objects	GILS record, by virtue of its creation, collects discrete information resources that record content indicates would not have otherwise been collected or aggregated. Assigned in the absence of clues within the record that the represented objects were heretofore packaged as <i>this collection</i> to optimize information discovery and retrieval.	<ul style="list-style-type: none"> • Privacy Act Systems compilation • files • press releases • forms
Aggregated Object Represented	GILS record represents an <i>a priori</i> or purposeful collection of information resources—e.g., woodpecker database or agency website. GILS record represents an object that collects, or comprises, two or more discrete information objects, and that represents a collection of standalone information files or products packaged together on the basis of a common theme or subject for functional convenience.	<ul style="list-style-type: none"> • CD-ROM of regulations • System that compiles Privacy Act records • job line of open requisitions
Discrete Object Represented	GILS record describes a standalone document-level entity that does not meet the criteria for “object aggregates metadata” below.	<ul style="list-style-type: none"> • annual report • videotape
Object Aggregates Metadata	GILS record describes a pre-existing metadata collection, or “locator,” as an information resource.	<ul style="list-style-type: none"> • directory • catalog • index • log

**Attachment E2-2a
Record Aggregates Objects**

Title: Briefing Materials, Public Comments, other Related
Official Files
Acronym:

Originator:
Department/Agency Name: Consumer Product Safety Commission

Local Subject Index:
Local Subject Term: US Federal GILS
Local Subject Term: Consumer Product Safety

Abstract: These files include the briefing memoranda prepared by the staff and reviewed by the Commissioners and public comments to any rulemaking documents.

Purpose: To adhere to the provisions of the Consumer Product Safety Act and other laws administered by the Commission.

Agency Program: To store, index and maintain the records associated with official Commission actions.

Availability:

Distributor Name: Office of the Secretary, Freedom of Information Division

Organization: Consumer Product Safety Commission
Street Address: 4330 East West Highway

City: Bethesda

State: MD

Country: USA

Zip Code: 20207

Telephone: 301-504-0800

Fax: 301-504-0127

Order Process: Official records are available in hard copy from the Office of the Secretary or the Office of Information and Public Affairs at the address provided..

Availability:

Distributor Name: Office of Information and Public Affairs

Organization: Consumer Product Safety Commission

Street Address: 4330 East West Highway

City: Bethesda

State: MD

Country: USA

Zip Code: 20207

Telephone: 301-504-0785

Fax: 301-504-0862

Order Process: Official records are available in hard copy from the Office of the Secretary or the Office of Information and Public Affairs at the address provided.

Access Constraints: None

Use Constraints: None

Point of Contact:

Name: Office of the Secretary
Organization: Consumer Product Safety Commission
Street Address: 4330 East West Highway
City: Bethesda
State: MD
Zip Code: 20207
Country: USA
Network Address: info@cpsc.gov
Hours of Service: 8:30 a.m. - 5:00 p.m.
Telephone: 301-504-0800
Fax: 301-504-0127

Control Identifier: CPSC0014

Record Source:

Department/Agency Name: Consumer Product Safety Commission

Date of Last Modification: 19960510

**Attachment E2-2b
Aggregated Object Represented**

TITLE: Worldwide Real Property Inventory System

ACRONYM: WRPIS

ORIGINATOR

DEPARTMENT/AGENCY NAME: General Services Administration (GSA)

MAJOR ORGANIZATIONAL SUBDIVISION: Public Building Service (PBS)

MINOR ORGANIZATIONAL SUBDIVISION: Office of the Chief Information Officer

NAME OF UNIT: Customer Service Division

ABSTRACT: The Worldwide Real Property Inventory System (WRPIS) supports the worldwide inventory information reporting cycle which involves collecting real property data (GSA FORM 1166) from Executive Branch Federal agencies, analyzing the data, and providing worldwide real property data to the public, private organizations, Congress, and other Federal agencies by providing a single source of information for both owned and leased property. The WRPIS system accepts data from Foundation Information for Real Property Management System (FIRM) as well as other media from the Executive Branch agencies. Reports and information available from WRPIS include the Summary Report of Real Property Leased by the United States Throughout the World, the Summary Report of Real Property Owned by the United States Throughout the World, and the World Wide Geographic Location Code Table.

BEGIN DATE: 1993

END DATE:

AVAILABILITY

DISTRIBUTOR

DISTRIBUTOR NAME: Public Buildings Service (PBS)

DISTRIBUTOR ORGANIZATION: General Services Administration (GSA)

DISTRIBUTOR STREET ADDRESS: 18th and F Streets, NW

DISTRIBUTOR CITY: Washington

DISTRIBUTOR STATE: DC

DISTRIBUTOR ZIP CODE: 20405

DISTRIBUTOR COUNTRY: USA

DISTRIBUTOR NETWORK ADDRESS:

DISTRIBUTOR HOURS OF SERVICE: 8:00 a.m. - 4:30 p.m.

DISTRIBUTOR TELEPHONE: 202-501-0856

DISTRIBUTOR FAX:

RESOURCE DESCRIPTION:

ORDER PROCESS:

ORDER INFORMATION: There is no on-line access to WRPIS outside of the Public Building Service except for the Worldwide Geographic Location Code Table. This information is accessible through the GSA Electronic Management Information (GEMI) Bulletin Board, 202-219-0132. This data is also available by printout and/or electronic media. To request printouts or electronic copies of the Worldwide Geographic Location Code Table or copies of the annual publications "Summary Report of Real Property Leased by the United States Throughout the World" and "Summary Report of Real Property Owned by the United States Throughout the World" write to the PBS Office of Governmentwide Real Property Policy (PG) at the address shown in the Distributor field.

COST:

COST INFORMATION: Some reports are free to other Federal agencies but there may be charges for the public and private sector. Specific costs will be determined on a case by case basis based on the specific requirements of the request.

TECHNICAL PREREQUISITES: To access GEMI Bulletin Board, 2,400 to 14,400 BAUD modem, ANSI or VT1000 terminal emulation, communications software

AVAILABLE TIME PERIOD

TIME PERIOD STRUCTURED:

TIME PERIOD TEXTUAL:

AVAILABLE LINKAGE: To access GEMI Bulletin Board, dial 202-219-0312, settings N-8-1-F.

AVAILABLE LINKAGE TYPE:

SOURCES OF DATA: Executive Branch agencies

ACCESS CONSTRAINTS

GENERAL ACCESS CONSTRAINTS: None.

DOCUMENTATION:

ORIGINATOR DISSEMINATOR CONTROL:

SECURITY CLASSIFICATION CONTROL:

USE CONSTRAINTS: None

POINT OF CONTACT FOR FURTHER INFORMATION

CONTACT NAME: Office of the Chief Information Officer

CONTACT ORGANIZATION: Public Building Service (PBS)

CONTACT STREET ADDRESS: 18th and F Streets, NW

CONTACT CITY: Washington

CONTACT STATE: DC

CONTACT ZIP CODE: 20405

CONTACT COUNTRY: USA

CONTACT NETWORK ADDRESS:

CONTACT HOURS OF SERVICE: 8:00 a.m. - 4:30 p.m.

CONTACT TELEPHONE: 202-501-9170

CONTACT FAX: 202-208-7087

PURPOSE: See abstract

AGENCY PROGRAM: WRPIIS supports the Federal Management Regulations 41 CFR 101-3 provision which requires Executive Branch Agencies, on an annual basis, to submit information regarding their properties to GSA.

SCHEDULE NUMBER: Not scheduled.

CONTROL IDENTIFIER: GSACP9002

RECORD SOURCE

RECORD SOURCE/DEPARTMENT AGENCY NAME: General Services Administration (GSA)

RECORD SOURCE/MAJOR ORGANIZATIONAL SUBDIVISION: Information Technology Service (ITS)

RECORD SOURCE/MINOR ORGANIZATIONAL SUBDIVISION: Office of GSA-Wide Information Technology (IT)

RECORD SOURCE/NAME OF UNIT: Center for GSA-Wide IT Systems Planning and Management

DATE OF LAST MODIFICATION: 19951031

SUPPLEMENTAL INFORMATION:

SYSTEM PRODUCTS DISPOSITION:

LOCAL SUBJECT INDEX

LOCAL SUBJECT TERM: US Federal GILS

**Attachment E2-2c
Discrete Object Represented**

Title: Investment Funds Brochure

Originator: U.S. International Development Cooperation Agency (IDCA)/Overseas Private Investment Corporation (OPIC)

Abstract: The Investment Funds Brochure contains general information on each of the Investment Funds guaranteed by OPIC. For each fund, the Fund Manager, Capital, Status of Funds Availability, Target Market, Targeted Sectors, Preferred Investment, Exit Strategy, Requirements, Qualifications for U.S. Business Participation, and Additional Services are listed.

Purpose: The Investment Funds Brochure is designed to provide a brief synopsis of OPIC's guaranteed funds to help American business executives and entrepreneurs interested in investing.

Agency Program: The Investment Funds Brochure describes Guaranteed Funds managed by the Investment Funds Program.

Availability:

Distributor Name: Information Officer
Organization: Overseas Private Investment Corporation
Street Address: 1100 New York Avenue, N.W.
City: Washington
State: D.C.
Country: U.S.A.
Zip Code: 20527-0001
Network Address: opic/s=info@mhs.attmail.com
Hours of Service: 8:45 a.m. to 5:30 p.m.
Telephone: 202/336-8799
Fax: 202/336-8700
Order Process: The Investment Funds Brochure is available without charge by writing to the Information Officer, at the above address, or by faxing a request to the number above, or by calling the Information Line at the phone number above.

Availability:

Distributor Name: Depository Library Program
Organization: Government Printing Office
Order Process: Item number OP.1.2; depository item number 0834-W-02.

Sources of Data: The Investment Funds Brochure provides information from sources within OPIC.

Access Constraints: None.

Use Constraints: None.

Point of Contact:

Distributor Name: Information Officer
Organization: Overseas Private Investment Corporation
Street Address: 1100 New York Avenue, N.W.
City: Washington

State: D.C.
Country: U.S.A.
Zip Code: 20527-0001
Network Address: opic/s=info@mhs.attmail.com
Hours of Service: 8:45 a.m. to 5:30 p.m.
Telephone: 202/336-8799
Fax: 202/336-8700

Schedule Number: Scheduled- N1-420-93-1, #26.

Control Identifier: IDCA/OPIC-GILS: 0008

Record Source:

Agency Name: U.S. International Development Cooperation Agency (IDCA)/Overseas Private Investment Corporation

(OPIC)

Major Organizational Subdivision: Management Services Department
Name of Unit: Information Center

Date of Last Modification: 19951218

335

**Attachment E2-2d
Object Aggregates Metadata**

Title: Office of the General Counsel Library Catalog

Originator: Pension Benefit Guaranty Corporation (PBGC)/Office of the General Counsel

Controlled Vocabulary: Yes

Abstract: The Office of the General Counsel Library Catalog describes the library's holdings. Approximately 1400 titles in the area of pensions and pension law, bankruptcy, administrative law, and Federal practice. The library is a Federal Government Depository and maintains a small collection of government legislation, regulatory, and other documents in print and electronic format.

Begin Date: 1990

Purpose: The Office of the General Counsel Library Catalog enables the user to locate library resources and materials. The catalog is used as an automated finding aid.

Agency Program: Library resources support the work of agency staff.

Distributor:

Name: Office of the General Counsel

Organization: Pension Benefit Guaranty Corporation

Street Address: Suite 340, 1200 K Street

City: Washington, D.C.

State: N/A

Zip: 20005-4026

Country: USA

Telephone: (202) 326-4004

FAX: (202) 326-4112

Order Process: Currently, there is no on-line access to the General Counsel Library Catalog. The catalog is available to users in the OGC library which is open to the public during business hours (below).

Available linkage: PBGC Home page URL: <http://www.pbgc.gov>
Available linkage type: html

Sources of Data: Inventory of library holdings collected internally from PBGC departments, and outside government agencies and sources.

Use Constraints: None

Point of Contact:

Lilian H. Fry, Librarian

Office of the General Counsel

Organization: Pension Benefit Guaranty Corporation

Street Address: Suite 340 1200 K Street

City: Washington, D.C.

State: N/A
Zip: 20005-4026
Country: USA
Hours: 9:00 a.m. - 4:00 p.m.
Telephone: (202) 326-4004
FAX: (202) 326-4112

Schedule Number: General Records Schedule Number #20.9.

Control Identifier: PBGC0003

Date of Last Modification: 19961119

337

APPENDIX E-3
Scripted Online User Assessment
Findings, Discussion, and Recommendations

APPENDIX E-3

Scripted Online User Assessment

Findings, Discussion, And Recommendations

Table of Contents

1.0.	Introduction	1
1.1.	Organization of Material	1
2.0.	Method Overview	1
3.0.	Data Summaries	2
3.1.	Participant Profile.....	2
3.2.	GILS Content Expectations.....	3
3.3.	GILS Service Expectations	4
3.4.	GILS Record Characteristics.....	5
3.5.	GILS Information Space and Ownership.....	5
3.6.	GILS Nomenclature	6
3.7.	Searching Reflexes and Relevance Judgments	6
4.0.	Findings, Discussion, and Recommendations	7
4.1.	Participant Profile Script Items	7
4.2.	Objectives-Driven Script Items	10
5.0.	Summary of Recommendations	28
5.1.	Increase Users' Searching Confidence	28
5.2.	Improve GILS Niche in Information Space.....	29
5.3.	Improve GILS Efficacy in NIDR and Revise NARA <i>Guidelines</i> Accordingly	29
5.4.	Improve the Quality and Consistency of GILS Records.....	29
6.0.	Areas for Further Research.....	29
7.0.	Conclusion	30
	Attachment E3-1: GILS Record from Script Item S28	31

APPENDIX E-3

SCRIPTED ONLINE USER ASSESSMENT

FINDINGS, DISCUSSION, AND RECOMMENDATIONS

1.0. INTRODUCTION

To capture user perceptions about and reactions to GILS concepts and implementations, the evaluation featured an exploratory technique based on a set of scripted service encounters. The scripted online user assessment achieved four goals. First, the online sessions permitted in-process, "front-line," collection of data concerning user assessments of GILS—as opposed to "recollection" of assessments after GILS use. Second, it elicited highly qualitative responses to a *concept* (i.e., rather than the more traditional aims of user assessments such as quantification of relevant "hits" or usage patterns). Third, the findings provide a degree of insight into the cognitive processes of users in the online, networked environment. Last, documentation of lessons learned during development and deployment of the new exploratory technique (see Appendix C-5 Scripted Online User Assessment Methodology) provides a basis future researchers to adapt the script and delivery techniques to their specific environments and objectives. Data such as those discussed in the following sections are crucial to understanding user perceptions, expectations, and behavior during networked information discovery and retrieval (NIDR), and in advancing the quality of GILS accordingly.

1.1. Organization of Material

Section 3.0 Data Summaries aggregates significant results in terms of the user session objectives. Detailed results of the sessions in Section 4.0 Findings, Discussion, and Recommendations are presented in order of appearance on the script and are prefaced by an alphanumeric code designating whether the data collection item (script question) was designed to profile the user (P) or support the specific objectives of the session (S). Each item in Section 3.0 Data Summaries tables cites the corresponding S or P number as found in Section 4.0 Findings, Discussion, and Recommendations. The codes "UNT" and SU" in the Section 3.0 data summary tables indicate that the finding resulted from the post-session debriefing at University of North Texas (UNT) or the post-session focus group at Syracuse University (SU).

Sections 5.0 and 6.0 present a summary of recommendations and opportunities for further research, respectively.

2.0. METHOD OVERVIEW

Graduate and undergraduate student "users" unfamiliar with GILS were oriented to the nature and purpose of their participation by means a 5-minute verbal introduction by the investigators. They were subsequently asked to record answers to more than 50 multiple-choice, free-form expression, and true/false questions as they navigated "real life," "real time" Government Printing Office (GPO) GPO Access and Environmental Protection Agency (EPA) GILS systems according to a scripted set of encounters. The script was based on results of the record content analysis (see Appendix E-2 Record Content Analysis Findings, Discussion, and Recommendations) and investigators' ongoing search/retrieval experience with various GILS. The questions were designed to elicit user feedback concerning GILS content and service expectations, record design, orientation in information space, adaptation to the metadata construct (e.g., searching reflexes), and, perhaps most importantly, users' assumptions about GILS—all on the basis of this 1-hour first-exposure to scripted transactions. In addition, investigators conducted debriefing sessions where users were informed generally of GILS scope and purpose and asked to elaborate on intellectual and emotional impressions created by the scripted. The qualitative data from the sessions were entered into a database to facilitate disclosure of patterns related to users' reactions to GILS as a service concept and to GILS product (search options, results set, and records). As with the record content analysis, investigators recorded suggested improvements to the development and execution techniques for scripted online-user assessment in order to optimize recommendations to agencies interested in adopting the techniques.

The scripted online user assessment script was developed and the sessions conducted during February 1997. Thus the results presented do no reflect any subsequent modifications to GPO's and EPA's GILS system configurations, capabilities, and user interfaces.

3.0. DATA SUMMARIES

The following tables summarize significant findings in terms of objectives for the user sessions. Investigators strongly recommend that interpretation of the following findings be guided by the complete data and discussions provided in Section 4.0 Findings, Discussion, and Recommendations, which are organized in script order, by an alphanumeric code designating whether the data collection item (script question) was designed to profile the user (P) or support the specific objectives of the session (S). Each Section 3.0 data summary table cites the corresponding S or P number as found in Section 4.0 Findings, Discussion, and Recommendations. The codes "UNT" and SU" in the Section 3.0 data summary tables indicate that the finding resulted from the post-session debriefing at University of North Texas (UNT) or the post-session focus group at Syracuse University (SU).

The scripted online user assessment script was developed and the sessions conducted during February 1997. Thus the results presented do no reflect any subsequent modifications to GPO's and EPA's GILS system configurations, capabilities, and user interfaces.

3.1. Participant Profile

The 10 participants represented reasonably capable but "GILS unaware" users of online networked information resources. The following table summarizes pertinent data.

Criteria / Findings Highlight	Source of Evidence/ Section 4.0 Item
Background	
1 "private citizen," 1 art undergraduate student, 1 history undergraduate student, 1 political science undergraduate student, and 6 library science graduate students	P5
Average of more than 2 years' Internet usage	P4
Government Information Experience	
Print sources of government information, on average, searched monthly or less frequently	P2a
Frequency of searching online sources of government information varies from weekly to "as required by class"	P2b
Reports on government activity/public notices and legislation most frequently sought information	P6
Only 1 participant had read, heard about, or used GILS (one encounter implied)	P3a, P3b
Most knew that Federal agencies have libraries	S13c*
Half of group unaware of the function or purpose of many Federal agencies	S32e*
Strong agreement that public electronic access to government information is important	S32k*
Searching Behavior	
Self-teaching through trial-and-error predominant method of acquiring/refining online searching skills	P1
Browsing websites or bookshelves more common than use of online help, card catalogs, or application of professional training	P7
User claims 40% "success" in locating government information by starting with agency homepages	UNT

*These (S) items, although appearing in the actual script portion of the instrument, reflect information about the user profile and thus are summarized here.

3.2. GILS Content Expectations

The following table summarizes session participants' expectations for GILS content—full-text of documents vs. metadata, subject matter and resource types, quality, scope and extent of collection, and record and resource aggregation (“distance” from satisfaction of an information need).

Criteria / Findings Highlight	Source of Evidence/ Section 4.0 Item
<i>Metadata Vs. Full Text of Documents</i>	
Abstracts; statements of where relevant information can be found and how to obtain it; and full-text of documents are predominant products expected from an (unidentified—i.e., theoretical and generic) “online information locator service”	P8
EPA GILS records describing a catalog of agency publications, technical reports, ozone statistics, and full-text of EPA regulations most expected	S26
Absence of full text (actual documents) causes “disappointment,” “surprise,” and “confusion”	SS9, SU, UNT
Given choice between a limited collection of full-text documents and a comprehensive collection of metadata, users prefer former	UNT
Agreement that GILS has enough fields to search	S32b, SU
<i>Subject Matter</i>	
Record describing a library was largely unexpected	S13d
EPA GILS records describing a catalog of agency publications, technical reports, ozone statistics, and full-text of EPA regulations most expected	S26
Availability of a free (no cost) document causes “interest”	S9
A toll-free number for ordering social security benefits expected	S20
“GILS is useful if you know what you’re looking for”	UNT
User attributes poor search result to ignorance of subject matter	S22
<i>Quality of Information</i>	
EPA information often expected to be the most current available	S30b
Availability of a free (no cost) document causes “interest”	S9
Field contents criticized as inappropriate (misused) and inadequate	S10, SU
User “frustrated” by record(s) brevity	S9
User finds “good, detailed information”	SU
A toll-free number for ordering social security benefits expected	S20
No consensus on whether all records should contain information in all elements	S32m
<i>Scope of Collection</i>	
Every agency publication not expected to be described in GILS	S30d
EPA GILS not assumed limited to headquarters information	S30g
Unclear how agencies choose what to include in GILS	S32j
Complaint of “long tedious lists” of hits	S22
User perceives GILS as “very comprehensive”	SU
Users want a “centralized federal server that integrates state information”	UNT
User perceives potential for world-wide GILS	UNT
<i>Granularity/Aggregation</i>	
Low consensus on information object described by EPA GILS record	S28
Varying record granularity perceived as a weakness	SU
“GILS is useful if you know what you’re looking for”	UNT

3.3. GILS Service Expectations

This table summarizes findings concerning session participants' expectations for GILS serviceability in NIDR. It includes their reactions to predictability of results, fielded searching, system errors and response time, hyperlinks, and GPO-EPA system consistency.

Criteria / Findings Highlight	Source of Evidence/ Section 4.0 Item
Predictability/Effectiveness	
Undirected subject-oriented search resulted in an average of 17 hits (max=40, min=2); quantities evoked "no surprise," "pleasantly surprise," "disappointment," and a sense of being "overwhelmed"	S6
"It's not the best search service out there"	UNT
"GILS is useful if you know it's there"	UNT
Fielded searching perceived as less than "helpful"	UNT
Logic and "Service Errors"	
Search engine logic failure produced "disappointment" and possible user abandonment	S9, S10, SU
Many causes perceived for "duplicate records"	S14
System/Service Speed	
"Rapid return of factual information" evoked "interest"	S9
Slow response rates frustrated users	SU
No consensus on GILS efficiency	S32h
"Comfortability"/Overall Satisfaction with GILS Concept and Design	
Agreement that GILS would be easier to search if records grouped hierarchically by subject	S32d
"Frustration" from "not knowing what to do with it [the record]" and "not knowing what [one is] looking at on the screen"	S9, SU, UNT
Future GILS usage prediction somewhat positive	S23g
Strong agreement that GILS is an improvement over microfiche and paper resources	S32l, SU
GILS providing "availability" to government information perceived as a strength, even if records are "non-pretty"	SU
User feels s/he's in a "trap" when searching GILS	SU
GILS assumes high degree of searcher sophistication; "[User] shouldn't have to feel like he's hacking into a government system" and "Would one turn a twelfth grader loose on GILS?"	SU, UNT
"Ideal/prototype user" of GILS seen as college student not "average citizen" (in terms of assumptions concerning information use)	UNT
GILS not considered "user friendly"	UNT
GILS has "potential"	SU
GILS has enough fields to search	S32b, SU
Hyperlinkage	
Hyperlink in "[Agency] Library Services" record title expected to lead to agency website or to OPAC	S11
Majority agree "all government documents should be hotlinked from one electronic card catalog"	S32c
Lack of hypertext criticized	S15
Implementation Policy/Consistency Across GILS(s)	
EPA GILS and GPO Access' GILS expected to operate "exactly the same" by some	S25b
EPA GILS not assumed to be mandated by law by most	S30c
EPA GILS assumed to be duplicated on GPO Access' GILS	S30e
Agreement GILS "probably helps agencies manage information resources"	S32i

3.4. GILS Record Characteristics

Study participants' reactions to characteristics of GILS records—e.g., cosmetic appearance, record length, element display order, and formatting are summarized below.

Criteria / Findings Highlight	Source of Evidence/ Section 4.0 Item
Cosmetics	
Lack of "pictures" criticized	S9
Existence of records, "even if non-pretty," seen as a strength	SU
Ergonomics	
Scrolling to get beyond index terms to "text" criticized	S9
"Flat, gray background" criticized as making "text harder to read"	S9
Length	
Record containing 14 elements perceived as "just right" length	S12
Record [space] "wasteful" relative to what it provides	UNT
File Formats	
Lack of HTML format criticized	S15
ASCII format errors criticized	S15, UNT
Format and completeness rank above accuracy and currency in evaluating records	S17
General	
Mild agreement "all GILS records should look alike"	S32f
Quality of records perceived to "vary widely"	S32n, SU
[Agency] Library Services record "satisfying" and "better" than that of another	S13b, S16

3.5. GILS Information Space and Ownership

The following table presents summary findings about study participants' perceptions of GILS navigability, centralization concepts (loci of services and products) and implementation architecture (network distribution), and availability and authority of resources.

Criteria / Findings Highlight	Source of Evidence/ Section 4.0 Item
Navigation	
"GILS is like a maze"	SU
Centralization	
[GPO Access' GILS] provision to search across agencies seen as strength	SU
EPA GILS thought to be "part of" GPO Access' GILS	S25c
Little surprise that EPA GILS "looks different" from GPO Access' GILS	S25a
GPO Access' GILS options perceived as publishers/distributors of information	S1
No consensus on probable number of GILS in existence or how to determine same	S31
Social Security database used to search for toll-free number for ordering social security benefits statement	S18a
Availability	
Web search engines believed to index GILS records	S30f
Authority	
Users assume that information in EPA GILS is authored by EPA	S30a

3.6. GILS Nomenclature

The following table summarizes participants' reactions to GILS presentation and use of bibliographical and NIDR terminology.

Criteria / Findings Highlight	Source of Evidence/ Section 4.0 Item
"Mandatory GILS core elements" expected to "always contain data"	S2
"Purpose," "Title," "Cross Reference," "Date of Last Modification" and "Sources of Data" highly ranked for "comfort/certainty of use" prior to searching	S3
Element definitions increased, decreased, and failed to affect "comfort/certainty of use"	S4
"Control Identifier" definition/role unclear; "document serial number" offered as alternative	S23,S24, UNT
Users "certain" of EPA GILS "Complete text," "Acronym," and "Local Subject Index" nomenclature prior to searching	S27
More than half misperceive "Date of Last Modification" as referring to resource rather than record	S29
Terms are "beyond comprehension of trained professionals"	SU
Elements are "misnamed" and "vague"	SU,UNT

3.7. Searching Reflexes and Relevance Judgments

Participants' preferences for full record vs. fielded searching, their relevance improvement tactics, and perceptions of user sophistication requirements (education and training) are summarized in the following table.

Criteria / Findings Highlight	Source of Evidence/ Section 4.0 Item
Fielded Searching	
Fielded searching perceived as less than "helpful"	UNT
GILS has enough fields to search	S32b,SU
"Purpose," "Title," "Cross Reference," "Date of Last Modification" and "Sources of Data" elements high for "comfort/certainty of use" prior to any GILS encounter	S3
"Local Subject Index" and "Controlled Vocabulary" elements among most popular for subject search	S5
Users "certain" of EPA GILS "Complete text," "Acronym," and "Local Subject Term" nomenclature prior to searching	S27
One-third tried fielded searching for known-item search	S18b
User attributes poor search results to ignorance of fielded search procedures	S22
Relevance	
Appearance of search term in record's title outrank "score" in evaluating hits	S8
Lack of precise recall (relevance of hits) criticized	S9, SU
Less than half judge a hit relevant from known-item search	S21
"Sophistication" Assumptions	
GILS assumes high degree of searcher sophistication; "[User] shouldn't have to feel like he's hacking into a government system" and "Would one turn a twelfth grader loose on GILS?"	SU, UNT
Most recognized appearance of named but theoretical search term in record	S13a
Social Security database used to search for toll-free number for ordering social security benefits statement	S18a
Various boolean expressions developed for known-item search	S18c
Only 1 user finds/recognizes/understands "Time Period of Content" element in record	S29
Agreement GILS is an improvement over microfiche and paper resources	S32l

4.0. FINDINGS, DISCUSSION, AND RECOMMENDATIONS

Detailed results of the sessions are presented below, in tabular form as appropriate, in order of the data collection item's appearance on the script. The prefatory alphanumeric code designates whether the data collection item (script question) was designed to profile the user (P) or support the specific objectives of the session (S). Throughout this section, the discussion of findings is correlated to results of the record content analysis (see Appendix E-2 Record Content Analysis Findings, Discussion, and Recommendations) as appropriate. In addition, recommendations based on interpretation of the findings are provided for improving GILS from a user perspective and for future research to clarify issues.

The scripted online user assessment script was developed and the sessions conducted during February 1997. Thus the results presented do no reflect any subsequent modifications to GPO's and EPA's GILS system configurations, capabilities, and user interfaces.

4.1. Participant Profile Script Items

The first 10 items plus 3 later in the script (S13c, S323, and S32k) captured demographic and other information about the participants—such as status (e.g., student, private citizen, etc.) and government-information search frequency, methods, resource types, and knowledge/attitudes about government information—providing a context for evaluating expectations and responses. Results of these items are summarized in the following section.

P1. How do you chiefly acquire or refine your online searching skills? (Circle one.)

Findings: 8 of 10 participants reported "self teaching through trial and error"; 1 participant reported "professional training" and 1 reported "reading online Help manuals." No one reported "applying knowledge of database design."

Discussion: Users, even library students, do not appear to rely on professional training in database design or searching methods for NIDR. Users may be reluctant to consult online help manuals to avoid interruptions in search-thought processes or having to interpret overly complex or technical-jargon-laden instructions. Also, online help might not be available in their experience, or, they may enjoy the challenge of "cracking" the system.

Recommendations: Present concise, comprehensible search instructions on the same page as the search input mechanism. Provide an example.

P2a. How often do you search *print* sources of government information?

Findings: 1 of 10 participants searches daily; one-third search monthly. The remainder search as required for academic credit and once or twice per year.

Discussion: Participants were not "power users" of printed government information. Adoption of access mechanisms for print sources to networked information resources may not bear fruit.

Recommendations: Additional research may be warranted to determine user satisfaction with agency name as the primary access point in traditional sources.

P2b. How often do you search *online* sources of government information?

Findings: 2 of 10 participants reported weekly searching and 2 reported monthly. The remainder search online as required for academic credits and up to 4 times per year.

Discussion: Overall, participants search online sources more frequently than print sources, although no one reported a frequency that in the investigators' judgment is required to gain retrieval proficiency.

Recommendations: Design GILS systems to accommodate the infrequent searcher.

P3a., P3b. Have you ever read or heard about the U.S. Government Information Locator Service (GILS), or actually used it?

Findings: Only 1 participant was aware of GILS: "found GILS when searching for class requirement."

Discussion: A small sample of users with backgrounds that one might expect to include a GILS encounter (see S5) were unaware of the service.

Recommendations: Advertise in government publications and libraries; incorporate GILS linkage into agency and White House homepages; register/index GILS homepages with popular Web search engines.

P4. Approximately how long have you used the Internet?

Findings: The maximum length of Internet usage was 4 years (1 participant); the minimum was 2 months (1 participant). On the average, users reported slightly more than 2 years' usage.

Discussion: Assumptions about potential GILS users in the academic environment (see S5) may reasonably include a relatively long Internet exposure/experience period.

Recommendations: Additional research is recommended to elucidate relationships between longevity of experience, self-ratings of "familiarity" or "proficiency," and satisfaction with system results.

P5. Please circle the letter that most closely matches your current status

Findings: Participants in the user study included 1 political science undergraduate student, 6 library science graduate students, 1 art undergraduate student, 1 history undergraduate student, and 1 "private citizen."

Discussion: The participants represented a reasonably "GILS-capable" population in terms of education level and subject orientation.

Recommendations: Additional user assessments should involve corporate librarians, small business owners, school teachers, political action group members (e.g., League of Women Voters), etc. It is suggested that Public Information and Freedom of Information Act Officers develop profiles of "print" information seekers, and that webmasters do likewise for agency website visitors, to optimize sampling for GILS user assessments.

P6. What types of government information do you seek most frequently? (Circle up to three).

Findings:

Type of Government Information Sought	N
Reports on govt activity/Public notices	7
Legislation	3
Research	2
Statistics	2
Budget and economic news	2
Case law	1
Historical	1
Regulations	1
International relations	1

Discussion: The seeking of information on government activities and for public notices may reflect a "news"-consumption behavior (i.e., a desire for "awareness" as opposed to specific, targeted information retrieval for "question answering") among students.

Recommendations: Additional research is recommended along the lines of "what was on your mind the last time you recall deliberately searching for or monitoring government information." For example, users seeking to satisfy a specific and direct but occasional information need may prefer the approach planned by USPS's WINGS (Web Interactive Network of Government Services) <<http://www.wings.usps.gov/Topten/>>, which will present information on, among other things, tax-return filing, requesting birth certificates, and job searching.

P7. Please circle letter(s) matching your experience seeking government information, in print or online
Findings:

Experience	N
I browse websites of bookshelves to find information	8
I have a few favorite sources that I have learned to use	3
I nearly always find just what I need	2
I begin my search with using a card catalog or online index	2
I avoid searching government sources directly whenever possible, and rely on secondary reports such as newspapers or CNN	2
I usually need help from a librarian or other intermediary to get started	1
I find that government information sources change often	1
I find user's instruction sorely lacking for most resources	0

Discussion: The experience of browsing was shared by the large majority of participants, as opposed to reliance on either a bibliographic tool (catalog or index) or human intermediary. No one reported experiencing a lack of user's instructions, which may indicate that they do not seek them out or that they find them adequate (see S1).

Recommendations: Additional research may find a relationship between "character of experience" and "type of government information sought" (see S6).

P8. If you were to enter search terms into an online "information locator service," what would you expect in return? (Circle all that apply).

Findings:

Online "Information Locator" Expectation	N
Abstracts or digests from relevant documents*	6
Statements about where relevant information is stored and how to obtain it*	5
Full text of documents that contain the information I seek	4
Relevant database names	4
Relevant document titles only*	3
A list of related, controlled subject terms from which to choose	3
A "frequently asked questions" (FAQ) list with answers	2
Names of experts in the subject	1
Other	1

Discussion: The expectations followed by an asterisk in the above table most closely match GILS. Bearing in mind that only 1 user had "heard or read about...or actually used GILS," it appears as though the participants had, prior to exposure, a fair idea of what to expect from the service. The notable exception is almost half of participants believing that GILS might provide full text of documents. It is interesting to note, however, that 3 of the 4 checking the "full-text" option also checked the "statements about where relevant information is stored and how to obtain it" option. Of the nine options presented in the script, the maximum N for any particular option was six and the minimum was one ("names of experts in the subject"). These data may reflect some users' notions about the varieties of product returned by searchable online resources and/or uncertainty about the terms "information" and "locator." Conversely, 1 participant's response indicated no ambivalence about service expectations: "exactly what I want."

Recommendations: Results of this and other items (see P9 and P260) support a recommendation for a more clear communication of GILS purpose and approach. The reader is referred, for example, to FedWorld's GILS site <<http://www.fedworld.gov/gils>>, which states directly above the search-form input boxes: "Please also note that GILS records are intended to allow you to learn about what government information is available, not to be [FedWorld's emphasis] the information that you might be seeking! [FedWorld's punctuation].". It is further recommended that agencies' avoid linking to or quoting verbatim policy documentation for the purpose of introducing users to GILS functionality.

4.2. Objectives-Driven Script Items

This section presents results of script items designed to support the research objectives—i.e., to capture user perceptions about and reactions to GILS concepts and implementations. (Note: The script was developed and the sessions conducted during February 1997. Thus the results presented do no reflect any subsequent modifications to GPO's and EPA's GILS system configurations, capabilities, and user interfaces.)

S1. There are many options listed underneath Individual Agency GILS databases on GPO Access. What do you think these might represent? (Circle all that apply).

Findings:

Databases Represent...	N
Publishers/distributors of information	7
Information creators	3
Internet server "mirror" locations	3
Other	"various govt agencies providing information to GILS"

Discussion: While it is difficult to interpret this finding in isolation from the efficacy of GPO Access GILS search page user-interface design, responses may indicate uncertainty about government information creation and dissemination policy and its implementation. Users with *a priori* knowledge of GPO's "distribution" mission may be especially confused; no explanation of the comparatively low incidence of "information creators" is offered. This issue is informed by S30a and S30e responses. The "Other" response also indicates some doubt or confusion; "various" is nonspecific, and the participant seems to imply that there is one GILS database to which agencies contribute.

Recommendations: GPO should include a straightforward statement on the GILS search page(s) to the effect that the databases contain an agency's GILS records of information resources created and available from that agency and that GPO GILS listing may be incomplete. In addition, the name "GILS" should be re-thought because it implies the singular. It may not be reasonable for a user to intuit that GPO's GILS is not the same as GPO Access' GILS; the former concept is comprehensible but the latter might be named more aptly "A Collection of Agency GILS." Likewise, since FedWorld's "GILS" doesn't exist in the singular it should not be so labeled on the website. All nonbrokered agencies' sites should be entitled "[Agency/Bureau/Etc. Name] GILS" rather than "GILS." S25a,b,c and S30e and S30a also address this issue.

S2. Would you assume that [GPO Access GILS] "mandatory GILS core elements" means that these fields always contain data?

Findings: 6 of 10 participants answered this question affirmatively.

Discussion: The term "mandatory" is misleading; its common synonym is "obligatory." Users familiar with commercial online search services, wherein "field" availability/existence is specified, may be especially prone to confusion.

Recommendations: The results of GILS record content analysis (see Appendix E-2 Record Content Analysis Findings, Discussion, and Recommendations) indicated an inconsequentially low incidence of "mandatory" element data population, even among those records designated as core by the presence of "US Federal GILS" in the local subject term element. This, coupled with the ambiguity of the words "mandatory" and "core," and perhaps even "element" (rather than "field"), should prompt an examination of the utility of exposing the user to the concept at all. S32m also speaks to this issue.

S3. Please rate how comfortable you would feel using the [GPO Access GILS] options presented under "Select one or more of the following fields...to search"

Findings: Prior to any searching and with instructions not to click on the hypertext to receive a definition, participants were most comfortable with Purpose, Title, Cross Reference, Date of Last Modification, and Sources of Data. They were least comfortable with, or most unsure of using, Original Control Identifier, Spatial Reference, Schedule Number, and Control Identifier.

Field	"Certain" N	"Unsure" N
Purpose	10	0
Title	10	0
Cross Reference	9	1
Date of Last Modification	9	1
Sources of Data	9	1
Availability	8	2
Local Subject Index	8	2
Point of Contact	8	2
Record Source	8	2
Time Period of Content	8	2
Access Constraints	7	3
Agency Program*	7	2
Controlled Vocabulary	7	3
Abstract	6	4
Originator	6	4
Supplemental Information	6	4
Methodology	4	6
Use Constraints	4	6
Control Identifier	2	8
Schedule Number	2	8
Spatial Reference	1	9
Original Control Identifier	0	10

1 participant did not address this option.

Discussion: Of common bibliographic metadata, participant confidence in the terms "Title," "Cross-Reference," and "Local Subject Index" was not unexpected. "Date of Last Modification" was revealed as problematic in the GILS record content analysis (see Appendix E-2 Record Content Analysis Findings, Discussion, and Recommendations) *vis-a-vis* misunderstanding that this element referred to the information resource being described rather than the record itself; the data above and that from S29 may corroborate this finding. The incidence (5) of "certain" understanding of or comfort using "Abstract" may be low, especially for the academically-oriented study group, and may reflect a distrust of the concept due to misappropriation of the term by many popular Web search engines.

Recommendations: Transaction log analysis may provide additional insight into users' choice of elements for fielded searching.

S4. [Participants were asked to read the GPO Access GILS-supplied definition of a field name they marked "unsure" of in S3] Does this definition affect your confidence in using this field for searching?

Findings: 7 of 10 participants noted that the selected definition "increased" confidence; 2 reported "no change" in confidence; and 1 reported a "decrease" in confidence.

Discussion: Nearly one-third of participants did not find the GILS element definitions helpful. S5, S14, S24, and S27 results also address fielded searching.

Recommendations: Subject the definitions to reading-level (e.g., Flesch Reading Ease, Flesch-Kincaid Grade Level, etc.) and grammar checks to compute "fog index" by identification of incomplete clauses, jargon, passive voice, characters/syllables per word and words per sentence.

These checkers are standard on most (fully-installed) popular word processing programs. Provide an example of the utility of the field to increase retrieval precision.

S5. Please check [of GPO Access GILS-supplied fielded searching options] below which ONE field you would select for locating information about native americans.

Findings: 4 of 10 participants selected Local Subject Index and 2 selected Controlled Vocabulary. The remaining 4 selected Title, Cross Reference, Agency Program, and Control Identifier.

Discussion: The reliance on controlled vocabulary for subject-driven searching as expressed by more than half of the participants may be a function of background or inclination; 3 of these 6 were library science students. There may also have been an assumption that these fields are always populated (see S2). The choices of Title and Agency Program may be interpreted as bids for high retrieval precision. However, the remaining choices appear unsupported by the current analysis, and may further indicate nomenclature problems (see S4).

Recommendations: In online help, provide an example of the utility of each field to increase retrieval precision.

S6. [As a result of executing a query by using All GILS Records on GPO Access GILS Site, the field selected in S5, and an (unrecorded) term "relevant to the concept of native americans"] How many total hits did you receive?

Findings: Of the 9 participants executing this search, the maximum report was 40 hits and the minimum was 2. Participants averaged 17 hits.

Discussion: The 2 participants receiving (default maximum) 40 hits searched the "Agency Program" and "Controlled Vocabulary" elements, respectively. 2 of the 3 participants receiving 2 hits (default failed-search Query Report and Database Catalog) searched the "Local Subject Index" element and the other searched "Control Identifier." A search on the "Cross Reference" element produced 32, "Title" produced 3, another's "Local Subject Index" search produced 4, and another's "Controlled Vocabulary" search produced 31 hits. Use of controlled vocabulary may have increased recall.

Recommendations: Survey the agencies using controlled vocabulary and determine, through log transaction analysis, whether the practice increases retrievals.

S7. [As a result of executing a query by using All GILS Records on GPO Access GILS Site, the field selected in S5, and an (unrecorded) term "relevant to the concept of native americans"] What is your reaction to the number of hits?

Findings: 2 participants did not answer this question.

Reaction To Total Number Of Hits	N
Not surprised	4
Pleasantly surprised	2
Disappointed but willing to examine the hits more closely	1
Overwhelmed but willing to examine some of the hits more closely	1
Disappointed but willing to start over with more specific search terms	0
Frustrated—I would abandon use of GILS at this point	0
Overwhelmed but willing to start over with more specific search terms	0

Discussion: Unfortunately, the participant expressing "pleasant surprise" received only 2 hits—the default "failed-search" results. Those "not surprised" received 40, 40, 4, and 2 hits, respectively, which may be indicative of a range of search-skill confidence levels. The "disappointed" participant received 3 hits, and the "overwhelmed" one received 32. It is interesting to note that no one expressed a desire to reformulate the query to produce fewer or greater hits. And, no one, with this first query in the script, was frustrated to the point of abandoning GILS.

Recommendation: With a more diverse and larger user sample, attempt to reproduce these results and correlate them with published findings of end-user search recall satisfaction (Note: no relevance judgment was required on this exercise).

S8. If you were to select one or more [GPO Access GILS-supplied] hits for closer examination, which factor would most likely influence your selection?

Findings: 1 participant did not answer this question.

Hit Selection Criteria	N
Appearance of search terms in the title	4
Score	3
Order of appearance (select first item first)	2
Format	0
Size	0
Other	0

Discussion: The preference for "appearance of search terms in the title" may reflect a lack of understanding or value of WAIS relevancy ranking. "Relevance is computed based on several factors, including the occurrence of the search terms in the document title, the frequency of the terms as a percentage of the total document size and conformance with the exact search phrasing."

(Ref: <http://www.access.gpo.gov/su_docs/aces/aces180.shtml?desc014.html#cont09>.

While WAIS will return highest "score" first, investigators included the "order of appearance" choice to acknowledge that several records may carry the same "score." It is interesting to note that no one assigned importance to file format (HTML, TEXT, PDF, SGML, etc.) or size (expressed in bytes).

Recommendation: None.

S9. Please characterize and explain your single FIRST REACTION to this record [retrieved on the basis of seeming "most promising" from the results list of the query on native americans at GPO Access' GILS]:

Findings: Some participants noted multiple "first" reactions.

Reaction	N	... Because ..
Disappointed	7	<ul style="list-style-type: none"> • The search found "american" in "anti-american" • I don't think it is what I really want • No pictures or nice formatting for user-friendliness • So flat, gray background made text harder to read • Was expecting to see the actual document • I didn't expect to get full text of "info" itself, but Abstract is too short and irrelevant • Not relevant
Surprised	3	<ul style="list-style-type: none"> • I got something • It isn't text—it's more of a list • I thought TEXT meant full text
Confused	2	<ul style="list-style-type: none"> • It did not give me a document but info on the document and where to write for a copy • I was looking at the subject index (investigators take this to mean "subject terms")—have to scroll down for text
Interested	2	<ul style="list-style-type: none"> • The page returned factual information rapidly • Tells me I can receive document free
Frustrated	2	<ul style="list-style-type: none"> • Not sure what to do with it • If this is best one, how brief are the others?

Discussion: This point in the script represented the first time 9 of 10 participants looked at a GILS record (see P3a). 7 reactions included expressions of "disappointment," and a combined 4 reactions were frustration or confusion. The most frequent comment concerned the lack of full text (4 participants); 2 expressing disappointment and 2 expressing surprise. (Both participants expressing "disappointment" over the lack of full-text also indicated in P8 that an online "information locator" might contain full text, but those expressing "surprise" did not!) 4 participants were disappointed by a perceived lack of relevance, and 2 participants were disappointed by their respective records' cosmetic appearance. Positive comments included "surprise" that a record appeared and "interest" evoked by rapid system response and availability of a resource at no cost. The comments of 2 users "frustrated" with their records are self-explanatory.

Recommendations: While the investigators believe that the options "surprised" and "interested" provided ample opportunity for positive feedback about participants' first impressions of a GILS record, it is recommended that a more direct record of user reactions be captured through "talk-aloud" protocol or a completely open-ended question.

S10. Please describe anything you consider to be peculiar or in error [about or in the retrieved record]:

Findings: Participants reported the following about their respective records (note that the script did not control which record users examined):

- Some of the descriptions do not appear to match their category (i.e., under agency program it describes how commission programs are designed to increase understanding—no actual info on agency program)
- Abstract should be abstract of the info, not only [just] description as to "report" and page nos.
- "Onondaga Nation" should retrieve hits with BOTH terms

Discussion: 2 of 4 participants responses to this item concerned unfulfilled expectations for content (in Agency Program and Abstract elements) and one concerned retrieval logic (see also S9). The record content analysis (see Appendix E-2 Record Content Analysis Findings, Discussion, and Recommendations) may corroborate incidence of element misuse as perceived by the study participants. An additional comment: "Five out of 12 hits were clicked and no data was returned. Could be time of day (12 pm)" was a result of poor wording of the question and was not considered in the analysis.

Recommendation: Agencies are encouraged to enlist objective content reviewers to evaluate conformance of fields' content to qualitative descriptions and examples provided *The government information locator service: Guidelines for the Preparation of GILS Core Entries* (National Archives and Record Administration, 1995a). In addition, search engine or search/retrieval-standard performance should be evaluated against system documentation (e.g., for GPO Access GILS [WAIS]):

<http://www.access.gpo.gov/su_docs/aces/aces180.shtml?desc014.html#cont01>.

S11. What do you think would happen if you were to click on this record's [retrieved deliberately by script instructions] hypertext title?

Findings: Note that the title of this record is "[Agency name] Library Services." 2 participants did not answer this question.

Result of Hypertext Title	N
I would jump to the [agency] website	4
I would connect to [agency] online library catalog	3
I would be given a list of library services such as interlibrary loan, photocopying, and research assistance	1
A list of [agency] library staff contact would appear	0
I would link to a fuller/longer version of this record	0
Other	0

Discussion: More than half of the responses correctly assumed a link to the agency's website. The remainder chose possibly plausible options.

Recommendation: In this particular instance, hypertext tagging of only the agency's name within the title of the record (i.e., Agency Name Library Services) might have resulted in a higher rate of correct answers. In all cases of hyperlinkage in GILS records, it is recommended that an objective party review the context of the link to ensure that the record creator is not over-assuming an intellectual "hop" that a user might not be prepared to make. The record content analysis (see Appendix E-2 Record Content Analysis Findings, Discussion, and Recommendations) also addresses this issue.

S12. Please rate your feeling about the length of this record [retrieved deliberately by script instructions] relative to your satisfaction:

Findings: Of the 8 participants responding to this item, 7 selected "Just right; it presented the necessary information" and 1 selected "Too long; it provided more than I needed to know. No one selected "Too short; it doesn't present enough detail."

Discussion: The subject record contained 14 elements and 4473 bytes—a relatively "short" record according to the record content analysis (see Appendix E-2 Record Content Analysis Findings, Discussion, and Recommendations). These limited data may indicate that users prefer to err on the side of brevity.

Recommendation: Additional user-based research could examine which elements contain the information (or information pieces) sought and/or which elements are consistently ignored by users.

S13a. Do you think the search term "trade agreements" would have produced a "hit" on this record [entitled "[Agency Name] Library Services" and retrieved deliberately by script instructions]?

Findings: Of 8 participants responding to this question, 6 correctly answered in the affirmative and 2 responded "no."

Discussion: The term "trade agreements" was one item in a bulletized list of subjects embedded in the record's Abstract element. This question was included to gauge participants' recognition or acknowledgment of search mechanics—i.e., in an indirect fashion determine whether, if a user were to search on "trade agreements" and receive a hit entitled "[Agency Name] Library Services" he or she would proceed to retrieve the record for further examination. In actuality, the subject database contains only 11 GILS records, and a search on "trade agreements" produces 2 hits (neither of which feature the term in the title; see S8), the second being "[Agency Name] Library Services." While it may be reasonable to assume that an agency library contains information on a wide variety of subjects, it may not be realistic to expect a user untrained in searching or unfamiliar with agency libraries to recognize this; the high recognition rate among the participants may be due to study group demographics (see P5).

Recommendation: Especially where records itemize subject areas covered, as in this case, the terms might be placed more appropriately in the Local Subject Index element. The GILS record content analysis (see Appendix E-2 Record Content Analysis Findings, Discussion, and Recommendations) discusses the issue of information resource granularity more thoroughly.

S13b. If you were looking for information about the [Agency Name] library, would this record [entitled "[Agency Name] Library Services"] satisfy you?

Findings: 7 of 8 participants were satisfied by this record.

Discussion: See S8 concerning users' evaluative criteria and S16 (relative quality rating).

Recommendations: "Model" records, as determined through user-satisfaction studies, should be readily available to serve as an example to record creators and as a benchmark for evaluators.

S13c. [Relative to record entitled "[Agency Name] Library Services"] Did you know that Federal agencies have libraries?

Findings: 8 participants answered this question; 6 with "yes"; 2 with "no."

Discussion: This question was included to bring perspective to S13d. Given the preponderance of library school students in our user sample, no conclusions are offered.

Recommendations: See S13d.

S13d. [Relative to record entitled "[Agency Name] Library Services"] Would you have expected to find a Government Information Locator Service record that describes a library?

Findings: 8 participants answered this question; more than half (5) answered affirmatively; 3 said "no."

Discussion: Interestingly, of the 5 "yes" respondents, 3 had reported "not expecting" an online information locator to provide "statements about where relevant information is stored and how to obtain it" (S8c).

Recommendation: Agencies with public service "traditional" libraries should cross-reference that resource within applicable records. The GILS record content analysis discusses the issue of information resource granularity more thoroughly (see Appendix E-2 Record Content Analysis Findings, Discussion, and Recommendations).

**S14. [Relative to a scripted retrieval of two duplicate records from a selected agency at GPO Access' GILS]
Why do you think two apparently identical results have been returned? (Circle all reasonable possibilities)**

Findings:

Reason for Duplication of Records	N
Both records describe the [agency name] library, but were created by different agencies	7
The titles of the two records have been shortened for this display; they are actually different	4
The system has made an error in searching or retrieval	2
The person who created the duplicate record was unaware that a record already existed	1
The search term [library] was too broad	1
The instructions I followed for this search are incorrect or incomplete	0
Other	"duplication of indexing"

Discussion: Note that the users were viewing a search-result page only at this point, not records themselves. The high incidence (7) of belief in two record sources may support a misunderstanding of database ownership and/or placement in information space (see also S1 S30a, and S30e) because participants executed this search in the "selected agency" mode rather than "all records on GPO Access GILS." The notion that titles had been truncated for display is certainly reasonable. It is interesting that only 2 of 10 users considered the possibility of system failure and that only 2 users attributed cause to human error (the searcher and the record creator, respectively). This result points to a tendency among participants to consider that "something about this thing called GILS" is at fault.

Recommendation: A mechanism such as input masks or prevent-duplicates indexing feature should be implemented by all GILS providers. The presence of duplicate records might easily erode user confidence in the quality of records or management of the system and can be easily avoided.

S15. Please describe anything you consider to be peculiar or in error [about this record retrieved according to the script]:

Findings: 8 of 10 participants responded to this question, and all but 1 of these noted the formatting error (apparently undelimited ASCII text; no hard returns):

- Wrap-around disabled; page info incomplete (avail)
- No HTML, only text was available
- No text wrapping—people like to scroll up and down, not left to right
- Page width is disconcerting and what kind of public personnel management is it. No additional link.
- Having to scroll to the right to read entire line
- That you have to scroll right and left to read is a problem
- Text runs horizontally only—must scroll to right to read text
- Have to expand screen to read full record, not enough information, poor information provision

In addition, 1 participant noted that the record "ended" prematurely (due to the formatting error) and 2 participants complained about the quality or incompleteness of the information itself. The comment concerning "no HTML" being available was assumed to be the result of a scripting error and was not considered in the analysis.

Discussion: Participants recognize formatting errors.

Recommendations: Record creators and/or approvers should view product as displayed by a browser selected on the basis of published usage reports (e.g., "Browser Battle" July 1996 *Internet World* p. 40) or their agency website access log analysis.

S16. Of the two agencies' records [scripted for retrieval] describing libraries, which is best?

Findings: Of 8 participants answering this question, 2 expressed "no preference"; the remainder thought the [agency name] library record was "better."

Discussion: The inadvertent omission of a question concerning "peculiarities and errors" vis-a-vis the preferred record precludes a definitive interpretation of this finding. However, it should be noted that the preferred record did contain hypertext, was correctly formatted and available in HTML, and contained discernable elements; the nonpreferred record did not feature these characteristics.

Recommendations: "Model" records, as determined through user-satisfaction studies, should be readily available to serve as an example to record creators and as a benchmark for evaluators.

S17. What characteristics distinguish the two records [describing agencies' libraries and scripted for retrieval] in your mind?

Findings:

Distinguishing Characteristic of Records	N
Format	7
Completeness	5
Accuracy	2
Currentness	2
Presence of hotlinks	2
Consistency*	0
Other	<ul style="list-style-type: none"> • I preferred abstract for ITC • ITC didn't work • comprehension of agency

* 2 participants did not address this option; 1 participant wrote "ignorable."

Discussion: These results are somewhat inconclusive absent operational definitions of the characteristics (e.g., "format" could be "accurate" or provide "hotlinks"), but may support the finding suggested in S9 that users place value on records' cosmetic appearance. It is interesting to note that 5 participants' responses imply that one or the other record was relatively "incomplete" when only 2 noted this in S15. The characteristics of "consistency" is further addressed by S32f, m, and n. In addition, it is noted that only one of the two records featured a hotlink and users were instructed (in the interest of time) not to pursue it; despite this limitation 2 participants recalled this as a distinguishing characteristic.

Recommendation: "Model" records, as determined through user-satisfaction studies, should be readily available to serve as an example to record creators and as a benchmark for evaluators.

S18a. Describe how you would use [GPO Access'] GILS to find the toll-free number for ordering a statement of earned social security benefits. What would you choose in the Make your [database] selection(s) scrollbox?

Findings: 9 of 10 participants completed this question. Of the 9, 8 reported that they would select "Social Security Administration," the ninth would select "all records on GPO Access site."

Discussion: (Note: the script was prepared and sessions conducted prior to the recent controversy surrounding availability of service via the agency's website.) This finding indicates that nearly all participants recognized the availability of a relevant and agency-specific database; the "all records" respondent may have been motivated by recall or perceived search efficiency rather than precision.

Recommendations: None.

S18b. Describe how you would use [GPO Access'] GILS to find the toll-free number for ordering a statement of earned social security benefits. Would you use "fielded search?"

Findings: Of the 9 participants answering the question, 3 would use fielded searching, the remainder would not.

Discussion: At this point in the script, participants had executed only one fielded search (S7-S10), for which the majority of reactions were negative (see S9). This initial turn-off may account for an apparent reluctance to use fielded searching even for this relatively specific information need.

Recommendation: It is recommended that GILS designers perform a confirmatory analysis that fielded searching improves retrieval precision in GILS for both known-item and exploratory searching.

S18c. Describe how you would use [GPO Access'] GILS to find the toll-free number for ordering a statement of earned social security benefits. What would you type into the "Enter your search term(s)" textbox?

Findings: 9 of 10 participants answered this question, providing the following search strings (presented here with the participant's method of acquiring online searching skills from S1):

Search String	Skills Acquisition
"benefits" AND "ordering"	self-teaching by trial and error
"earned benefits"	professional training
earned benefits	self-teaching by trial and error
earned benefits	self-teaching by trial and error
"social security" AND benefits	self-teaching by trial and error
social security benefit*	reading online Help manuals
"toll-free number" ordering benefits	self-teaching by trial and error
toll-free number	self-teaching by trial and error
toll-free social security phone number	self-teaching by trial and error

Discussion: The script instructed participants to read a brief and accurate description of boolean operators in preparation for a previous searching event. Comparable information was present on their screen during the present exercise (<http://www.access.gpo.gov/su_docs/gils/gils.html>). Of the 8 participants who selected the social security administration database (S10a), all but 2 omitted the term "social security" from their search string; the participant who opted to search "all records" included the term and reported use of online help to acquire or refine searching skill (S1). It is interesting to note that only 2 users used the term "ordering," which is the operative "action" or "service" word in this concept. In addition, the three instances of "toll-free" may indicate an attempt at inappropriately high precision in specifying the type of information resource (i.e., only 1 user included the less-precise term "phone", which would not retrieve "telephone").

Recommendations: None.

S19. [At GPO Access' GILS and as a result of executing an unscripted query concerning a "toll-free number for ordering a statement of earned social security benefits"] How many "total hits" did GILS return?]

Findings: 9 participants answered this question. Results are shown below next to search terms used.

N Hits	Search String
40	earned benefits
40	toll-free social security phone number
14	"social security" AND benefits
14	"toll-free number" ordering benefits
3	"benefits" AND "ordering"
3	"earned benefits"
3	earned benefits
3	toll-free number
2	social security benefit*

Discussion: By way of background and context, this nonscripted (i.e., user-directed) search exercise was included on the basis of an investigator's recall of having used such a service more than 5 years ago. The investigators attempted several search strings within SSA's GILS to locate the telephone number prior to the user session, and then replicated study participants' queries using the above data—all to no avail. Acting on an impression that the benefits-statement service had been quite popular and may still be available, the investigator visited the USPS's WINGS (Web Interactive Network of Government Services) <<http://www.wings.usps.gov/Topten/>> website, which is designed to provide direct access to frequently requested "pieces" of government information. Under WINGS' "retirement" category one finds hypertext "Social Security Benefits/How do I...," which links to SSA's "Personal Earnings and Benefit Estimate Statement" (PEBES) website <<http://s00dace.ssa.gov/pro/pebes/pebes-home.shtml>>. At this site, a user can request (via forms interface or email but apparently no longer by telephone) the subject benefits statement. When the "official" term Personal Earnings and Benefit Estimate Statement (no quote marks)

was entered into SSA's GILS, however, only one record was retrieved. This GILS record, entitled "Earnings Record And Self-Employment Income System," contains in the Purpose: "Master machine-readable file containing a summary of earnings for all individuals, including the self-employed, who pay social security taxes." However, this US Federal GILS "core" record also states that Availability is "none." (In addition, Record Source contains the word "none.") Unfortunately, a search of SSA's GILS (which contains 1203 records according to the WAIS catalog) by use of "web* OR URL OR home* OR Internet OR http*" also failed to produce a record for the agency's website, which may have led a user to the PEBES feature.

Having confirmed the correct name for the PEBES service, the investigator analyzed the possibility that any of our study participants *could have* retrieved a GILS record if one existed by using their search strings. Only "social security benefit*" would have theoretically retrieved the record. Investigators' attempt to replicate this user's search (against "All records on GPO Access site," however, returned the default-maximum 40 hits, indicating that the user either failed to record his search strategy correctly or introduced a typographical error. The remainder of the search strings would have failed variously due to lack of truncation ("benefits" will not retrieve "benefit"), use of constraining quote marks, or use of terms based on an assumption of a telephonic rather than digital transmission medium.

Recommendation: See S22.

S20. [At GPO Access' GILS and as a result of executing an unscripted query concerning a "toll-free number for ordering a statement of earned social security benefits" Did you expect a "hit" that would obviously point you to the toll-free number?]

Findings: 4 of 9 participants providing an answer to this question responded "yes"; the remainder did not expect a relevant hit.

Discussion: See S22.

Recommendations: See S22.

S21. [At GPO Access' GILS and as a result of executing an unscripted query concerning a "toll-free number for ordering a statement of earned social security benefits"] Do any "hits" appear to be relevant?

Findings: 4 of 9 participants providing an answer to this question responded "yes"; the remainder did not discern a relevant hit.

Discussion: Interestingly, only 2 of the 9 users had their "expectation of success or failure" fulfilled (i.e., in 7 cases S20 results mapped inversely to S21 results). See also S7 and S22.

Recommendations: See S22.

S22. [At GPO Access' GILS and as a result of executing an unscripted query concerning a "toll-free number for ordering a statement of earned social security benefits," if you did not expect a relevant hit or if you did not receive a relevant hit] Why not? (Circle all reasonable possibilities.)

Findings: 9 of 10 participants answered this question.

Reason for Search "Failure"	N
I don't know enough about social security to come up with good search terms or to choose database(s)	4
I'm unsure of "fielded searching" in GILS	2
The toll-free number is probably too new to be included in GILS	2
I doubt if such a service exists	1
The phone number is probably on a website and therefore not duplicated in GILS due to the maintenance burden	1
I don't think GILS would include telephonic information resources	0
The Social Security Administration does not participate in GILS	0
Other:	
• I usually get long tedious lists of unnecessary links	
• I neglected to ask for "toll-free number"	2

Discussion: Refer to S18 for an analysis of users' search terms; "choice of database(s)" is moot given that all participants used either "All records on GPO Access site" or "Social Security Administration." Interestingly, of the 2 participants feeling "unsure of fielded searching," only one actually utilized that feature in the exercise. The options of (probably) "too new," "doesn't exist," and "on a website" may be indicative of some uncertainty about

policy and procedures for GILS content and its maintenance; S32j speaks to this issue as well. The investigators interpret the "long tedious list" comment to mean that the user did not wish to evaluate returned hits for relevancy (this user did not "expect" (S20) but did report (S21) at least one promising hit of the 40 produced). The user who "neglected to ask for 'toll-free' number" used "earned benefits" as a search string in the SSA GILS and received 3 hits (default "failed search" hits). The belief that the telephone number might be on a website ($N=1$) proved plausible (see S19).

Recommendations: Additional research into users' attribution of error/failure cause will inform development and continuous improvement of online help facilities.

S23. [After reading the GPO Access GILS-supplied definition of "control identifier"] Is the definition, and how it fits into GILS, clear to you?

Findings: 8 of 10 participants answered this question; 3 answered "yes" and 4 answered "no."

Discussion: The supplied definition was: "This element is defined by the information provider and is used to distinguish this locator record from all other GILS Core locator records. The control identifier should be distinguished with the record source agency acronym provided in the U.S. Government Manual." This definition is as published (in part) in the NARA *Guidelines*. While the present study did not ask users to try to pinpoint the source of confusion, we might assume the following. The word "element" means less to users than to GILS record creators, and users encounter this definition by pursuing links that refer to "fields" rather than elements.

"Information provider" is also a vague term (e.g., a user could conceivably assume that it refers to database "owner," record creator, record source, the U.S. Public Printer, the agency's public information officer, etc.). The definition might incorrectly assume an understanding or appreciation of the concept of "core" and "locator" records. Finally, users familiar with the contents of the *U.S. Government Manual* may be confused by the definition's implication that the publication actually provides a "record source agency acronym." (S4 also addresses GILS nomenclature.)

Recommendations: See S4.

S24. [Upon searching for a predetermined-as-duplicate control identifier at GPO Access' GILS] Keeping in mind that you have searched the "control identifier" field, whose contents "distinguish this locator record from all other GILS Core locator records," what is your reaction to the list of hits?

Findings: 3 participants did not answer this question.

Reaction to Duplicate Control Identifier	N
I do not understand "control identifiers"	4
One of these is something other than a GILS Core locator record	1
The record creators made an error	1
The records are the same—one in English and one in Spanish	1
I do not notice anything unusual about this search result	0
The system has made an error	0
Other	0

Discussion: Of the 5 users who responded negatively to S23 (definition was not clear to them), the 3 who also answered this question (S24) affirmed this confusion. The participants who attributed duplication to "core locator record" status and record-creator error both answered S23 affirmatively (definition was clear). The participant who selected the multilingual explanation responded to S23 negatively (definition was not clear). It is interesting to note that no users assumed system error.

Recommendations: If a purpose of control identifiers is to uniquely identify all records contained in all GILS (with the result of absolutely precise retrieval in known-item searching), a mechanism such as input masks or prevent-duplicates indexing feature should be implemented by all GILS providers.

S25a. [As a result of linking via browser bookmark to "U.S. Environmental Protection Agency—Search the GILS Database] Are you surprised to find that EPA's Government Information Locator Service looks different from GPO Access' Government Information Locator Service?

Findings: 8 of 10 participants answered this question. 3 responded affirmatively and 5 negatively.

Discussion: One explanation for "surprise" might lie in the title of this page, which could be interpreted to mean something other than an agency-specific GILS (i.e., a user, especially if under the impression that there is only one integrated GILS database, could conceivably believe that this site is EPA's rendition of GPO Access' GILS

resource). (See also S1, S18a, and S31). "Lack of surprise" might be attributable to websurfers expectation of variation (e.g., among agency homepages or among the .edu, .gov., and .org servers of the same government legislative databases).

Recommendation: All nonbrokered agencies' sites should be entitled "[Agency/Bureau/Etc. Name] GILS" rather than "GILS" or in some manner make apparent that they are a subset of the GILS universe.

S25b. [As a result of linking via browser bookmark to "U.S. Environmental Protection Agency—Search the GILS Database] Do you expect the EPA's Government Information Locator Service to operate exactly like GPO Access' Government Information Locator Service?

Findings: Of the 8 participants responding, they were evenly divided between "yes" and "no."

Discussion: This finding perhaps corroborates the interpretation for S25a in that users appeared to manifesting some confusion about standard operability and/or agency leeway. (EPA's GILS, in fact, does not "operate exactly like GPO Access' GILS"—for example it, does not return search terms with results, it contains a "browse" feature, and offers a different set of elements for fielded searching.)

Recommendation: More extensive cross-GILS research could reveal whether diversity in presentation/operability approaches is a strength or weakness for end-users.

S25c. [As a result of linking via browser bookmark to "U.S. Environmental Protection Agency—Search the GILS Database] Do you think that EPA's Government Information Locator Service is part of GPO Access' Government Information Locator Service?

Findings: Of the 8 users answering this question, 5 answered affirmatively and 3 negatively.

Discussion: The term "part of" might have been construed as "included on" the GPO GILS site (which it is not), or as "in cooperation with" GPO. Users also might have been under the impression that GPO is "in charge" of the GILS initiative and interpreted "part of" to mean "under the aegis of."

Recommendations: See S1. Also, popular government-information-seeking starting points (as determined by log transaction analysis, but we may assume the White House homepage, Library of Congress homepage, and GPO Access for examples) should consider featuring a link to an information-space map of GILS that shows host overlaps and organizational relationships.

S26. [As a result of linking via browser bookmark to "U.S. Environmental Protection Agency—Search the GILS Database] What might you expect EPA's Government Information Locator Service to provide? (Circle all that apply).

Findings: (8 of 10 participants responded to this question.)

EPA GILS Content Expectation	N
A catalog of EPA publications	8
Descriptions of technical reports	7
Statistics about the ozone layer	6
The full text of EPA regulations	5
Congressional testimony on nuclear accidents	5
Hotline phone number	5
A phone directory of EPA staff	5
A list of Superfund cleanup sites	4
An order form for a radon-testing kit	3
Census data	3
Images of the spotted owl	3
Hotlinks to environmental activism websites	3
Maps	2
An abstract for a CD-ROM about nature	2
Clinton's 1996 inaugural address	1

Discussion: At this point in the script, participants had executed five searches at GPO Access' GILS: three against "All records" and two against specific agencies' databases. Two of the five searches were fielded and the remainder "full-text." These searches resulted in users' examination of a minimum of three GILS records.

The overall result of this question indicates some degree of understanding GILS to be a "locator of locators" (all participants expected a publications catalog). However, about two-thirds of the users also expressed expectation

of actual documents (the "information itself") such as regulations and testimony *even though they had no precedence for this belief in practice*. Interestingly, only 1 of the 5 users expecting "full-text of EPA regulations" had expected an "online information locator" to contain "full-text of documents containing the information I seek" in P8. In addition, the notion that EPA GILS might provide census data or an inaugural address indicate confusion about the agency-specific content of each GILS database and may be related to information-space disorientation: 3 of the 5 users who thought EPA GILS was "part of" GPO Access' GILS (S26c) expected to find census data on EPA GILS. In summary, with this limited line of inquiry, we may conclude that *users reluctantly abandon an expectation of direct access to actual documents/resources and/or do not adopt readily to the GILS system of metadata records*. **Recommendation:** Caveats about content (such as that provided on FedWorld's GILS site <<http://www.fedworld.gov/gils>>: "Please also note that GILS records are intended to allow you to learn about what government information is available, not to be the information that you might be seeking!") appear to be warranted.

S27. Please rate how confident you would feel using the following [EPA GILS fielded search] options presented on this screen:

Findings: 8 of 10 participants answered this question.

EPA GILS Search Option	"Certain" N	"Unsure" N
Complete text	8	0
Acronym	5	3
Local Subject Term	5	3
Agency Program	4	4

Discussion: Of the options presented, Local Subject Term and Agency Program appear in item S3 (similar assessment of terms at GPO Access' GILS), but users' assessment of confidence declined during the intervening GILS experience. It is noted as well that, while S3 did not assess this option, GPO Access' GILS equivalent phraseology for EPA's GILS "complete text" is *full text*. The word choices for both systems appears to be risky in light of users' expectation to access actual "information itself" (S8 and S26).

Recommendations: In addition to those provided in S8 and S26, it is suggested that the terms such as "complete text" and "full text" as search options be replaced with "all fields."

S28. [As a result of scripted (directed) retrieval and "scanning" of an EPA GILS record entitled "Index to the Wetland Educational Resources distributed by the New England regional office, EPA"] Which of the following does this record describe?

Findings: 7 of 10 participants answered this question; 1 participant indicated 2 responses.

EPA GILS "Resource" Description	N
An index	3
An educational "kit"	1
Miscellaneous training items available separately	1
Don't know	0
Other:	2
• a federal regulation project to protect wetlands	
• like an infoguide	

Discussion: For context, the subject record is reproduced in Attachment E3-1. The GILS record content analysis (see Appendix E-2 Record Content Analysis Findings, Discussion, and Recommendations) addresses resource description and information object identification in detail. However, the results of the present assessment indicate that record titles are powerful and that aggregation of resources within a single GILS record may be problematic. It is believed that the subject record describes "miscellaneous training items available separately." The concept of a "kit" however, may be inferred from the record's statement that "Supplemental Information: Information collection [investigators' emphasis] has a particular emphasis on wetlands stewardship materials for educators teaching grades K-12." Further, the "other" comments by these users corroborate qualitative results from the record content analysis indicating that "agency program," "purpose," "information resource" elements' content are not consistently distinguishable in GILS records (see Appendix E-2 Record Content Analysis Findings, Discussion, and Recommendations). It is possible that the participant who suggested "like an infoguide" inadvertently described a *result* of aggregation rather than the actual materials.

Recommendations: The Resource Description element should be mandatory and its content be drawn from a controlled thesaurus. (See Appendix E-2 Record Content Analysis Findings, Discussion, and Recommendations for one approach to operationalizing information object/container terminology.)

S29. How up-to-date are the described materials [described in the EPA record] ?

Findings: The following responses were received from this fill-in-the-blank question:

- December 6, 1995
- 1995 December 6
- 12/6/96 — over 1 yr old
- from 1995 I think
- don't know
- hard to tell from site
- not supplied

Discussion: About half of the participants answering this question referred to the record's Date of Last Modification element (see Attachment E3-1)—a mistake common with GILS records creator according to the GILS record content analysis (see Appendix E-2 Record Content Analysis Findings, Discussion, and Recommendations). One of these participants offered what might be a value judgment in addition to the requested answer: "over one year old." In fact the record states "Time Period of Content: not supplied," which *does* refer to the information resource itself.

Investigators assume that the participant who responded "hard to tell from site" (interpreted as "record") might have expected a "date of publication" field (which investigators' acknowledge as not equivalent).

Recommendations: Adopt more precise nomenclature for Data of Last Modification (e.g., "GILS Record Revision Date" and "Time Period of Content" (e.g., "Time Span of Featured Subject Matter"). In addition, it is recommended that "publication dates" be required as available, and that historical resources cross-reference current ones as available.

S30a. [After scripted retrieval and examination of a record from EPA GILS search page] Would you assume true/false: Information in EPA GILS is authored by EPA

Findings: 7 of 10 participants answered this question. 5 answered "true," the remainder "false."

Discussion: Some users are uncertain of "database owner," "URL host," "GILS Provider," and other responsibility and authority boundaries.

Recommendation: See S1.

S30b. [After scripted retrieval and examination of a record from EPA GILS search page] Would you assume true/false: Information in EPA GILS is the most current available

Findings: 7 of 10 participants answered this question. 4 answered "true," the remainder "false."

Discussion: Users do not universally assume that information on the Internet is current.

Recommendation: See S29.

S30c. [After scripted retrieval and examination of a record from EPA GILS search page] Would you assume true/false: EPA is mandated by law to provide the information in GILS

Findings: 7 of 10 participants answered this question. 5 responded "false," the remainder "true."

Discussion: Study participants were not exposed to any form of GILS policy documentation during the session, and the 1 participant having experienced a previous GILS encounter (P3b) did not provide a response to this question.

On this basis, investigators conclude that these responses represent guesses, and no conclusions may be drawn.

Recommendations: It would be interesting to pursue whether users informed of GILS purposes and mandate assess GILS differently—i.e., whether inconsistencies and errors are less tolerated.

S30d. [After scripted retrieval and examination of a record from EPA GILS search page] Would you assume true/false: EPA GILS describes every EPA publication

Findings: 7 of 10 participants answered this question. Only 1 participant answered "true," the others "false."

Discussion: Given that the scripted activities provided no direct indication of the number of records in this agency's database, the finding indicates that users may appreciate the unfeasibility of describing all publications, much less "information resources" in a GILS database.

Recommendations: The concept of "Core" records should be re-specified or abandoned, as it is the sole yardstick for grasping the extent of Federal information resource holdings described in GILS. In addition, agencies should state on their GILS site what criteria how resources are chosen for description by a GILS record (S30j also speaks to this issue).

S30e. [After scripted retrieval and examination of a record from EPA GILS search page] Would you assume true/false: A duplicate of EPA GILS exists on GPO Access' GILS

Findings: 7 of 10 participants answered this question. 5 responded "true," and the remainder "false."

Discussion: The participants had examined GPO Access' GILS scrollbox of participating agencies to answer S1 and used the scrollbox in executing various searches. The EPA GILS page they encountered per the script did not refer to GPO Access' GILS. In light of these observations, we could conclude that users may assume (perhaps by virtue of recalling an option for "All records" in the GPO Access GILS scrollbox) that all agencies' GILS are served by GPO. (S31 addresses this perception more directly).

Recommendation: See S1.

S30f. [After scripted retrieval and examination of a record from EPA GILS search page] Would you assume true/false: I can find EPA GILS records by use of a web search engine such as Yahoo!, Alta Vista, or Lycos

Findings: 7 of 10 participants answered this question. 5 responded "true," the remainder "false."

Discussion: Users may assume that Web search engines automatically index all Internet content or that GILS database providers/owners have "registered" their content with popular search engines.

Recommendation: In GILS marketing, make this a "plus" (i.e., state that *only* through use of GILS can users directly access descriptions of thousands of agency resources). In addition, GILS homepages should be registered with popular Web search engines.

S30g. [After scripted retrieval and examination of a record from EPA GILS search page] Would you assume that this GILS has only information resources of EPA Headquarters in Washington, and does not include regional offices

Findings: 7 of 10 participants answered this question. 5 participants believed this statement to be "false," 2 believed it to be "true."

Discussion: The question was included as a measure of participants' recall that the title and other elements of the retrieved record stated clearly that the EPA information resource was "regional." Some users may require a more direct disclaimer as to the organizational scope of GILS records.

Recommendations: GILS sites should state the scope of the records collection in terms of, among other characteristics, organizational boundaries.

31. [After approximately 1 hours' GILS experience comprising 5 searches of GPO Access' GILS (both "All records" and selected agency database(s) options) and 1 search on EPA GILS] How many GILS do you think may exist on the Internet today? (Circle one.)

Findings: 9 of 10 participants answered this question.

Number of GILS in Existence	N
One for each Federal agency	3
I have no basis for guessing	2
One for each website that has ".gov" as part of the URL address	2
One for each branch of the government	1
One for each type of information resource	1
One for each broad subject area	0
Only one	0

Discussion: After approximately 1 hours' GILS experience comprising 5 searches of GPO Access' GILS (both "All records" and selected agency database(s) options) and 1 search on an independent (nonbrokered) GILS, the finding that two-thirds of participants selected option other than "one for each Federal agency" indicates uncertainty about GILS scope, placement in information space, and/or "ownership."

Recommendations: See S1. The script assumed that information seekers are generally not motivated to link to and digest documentation concerning system or service policy. It was noted during development of the user session and

record content analysis that a high number of agencies have mounted OMB 95-01 and other policy documentation, presumably by way of *informing* visitors about GILS rationale. This practice should be supplemented by an *educating* document that contains a standard (i.e., government-wide) statement about the GILS universe and the host's placement within it.

S32a. [After 1 hour's experience comprising 6 scripted searches on GPO Access' and EPA's GILS plus up to 10 minutes nonscripted exploration of up to 5 nonbrokered GILS]

All agencies' GILS should be searchable together, from one website.

Findings: 9 of 10 participants answered this question.

Strongly Agree	Agree	Disagree	Strongly Disagree	No Opinion
5	2	1	0	1

Discussion: Centralization of access is a desired state.

Recommendations: None.

S32b. [After 1 hour's experience comprising 6 scripted searches on GPO Access' and EPA's GILS plus up to 10 minutes nonscripted exploration of up to 5 nonbrokered GILS]

There are not enough fields to search in a GILS database.

Findings: 9 of 10 participants answered this question.

Strongly Agree	Agree	Disagree	Strongly Disagree	No Opinion
0	1	5	1	2

Discussion: Of participants expressing an opinion, there are a sufficient number of fields to search.

Recommendations: Assess fielded-search usage in actual practice by means of log transaction analysis.

S32c. [After 1 hour's experience comprising 6 scripted searches on GPO Access' and EPA's GILS plus up to 10 minutes nonscripted exploration of up to 5 nonbrokered GILS]

All government documents on the Internet should be hotlinked from one electronic card catalog.

Findings: 9 of 10 participants answered this question.

Strongly Agree	Agree	Disagree	Strongly Disagree	No Opinion
1	5	1	0	2

Discussion: This question assumes "full document retrieval" or "access to the actual information." Respondents by a reasonable margin prefer "one-stop shopping."

Recommendations: Implementation of GILS on Z39.50-compliant servers and increased description of online resources will promote a perception of "seamless" service for GILS.

S32d. [After 1 hour's experience comprising 6 scripted searches on GPO Access' and EPA's GILS plus up to 10 minutes nonscripted exploration of up to 5 nonbrokered GILS]

It would be easier to search GILS records if they were grouped hierarchically by subject.

Findings: 9 of 10 participants answered this question.

Strongly Agree	Agree	Disagree	Strongly Disagree	No Opinion
0	6	1	0	2

Discussion: This result of two-thirds "agreeing" to the serviceability of subject-oriented access is interesting. GILS is based on the GPO model of agency name as primary access, as is most of the White House website. More than half of our well-educated demographic (P5) indicating a lack of awareness of agency functions (S32e) points to the need for alternative approaches to locating government information.

Recommendations: Further research appears warranted in two areas: the incidence of subject-oriented inquiry vis-a-vis GILS and the feasibility of a non-organization-based classification scheme. Also see S19.

S32e. [After 1 hour's experience comprising 6 scripted searches on GPO Access' and EPA's GILS plus up to 10 minutes nonscripted exploration of up to 5 nonbrokered GILS]

I am unaware of the function or purpose of many Federal agencies.

Findings: 9 of 10 participants answered this question.

Strongly Agree	Agree	Disagree	Strongly Disagree	No Opinion
2	3	4	0	0

Discussion: From this nearly evenly divided result from a well-educated demographic (P5), we may conclude that GILS implicit assumption of agency-mission knowledge is unwarranted. See S32d.

Recommendations: Agency GILS "index.htm" or search pages could provide a prominent link to their mission statement and/or a list of general functions (perhaps from *the U.S. Government Manual*).

S32f. [After 1 hour's experience comprising 6 scripted searches on GPO Access' and EPA's GILS plus up to 10 minutes nonscripted exploration of up to 5 nonbrokered GILS]

All GILS records should look alike.

Findings: 9 of 10 participants answered this question.

Strongly Agree	Agree	Disagree	Strongly Disagree	No Opinion
1	4	2	1	1

Discussion: Participants expressed a "soft" preference for consistency in the appearance of GILS records.

Recommendations: Further research could operationalize "look alike"—e.g., determine whether this preference considers file format (HTML, PDF, ASCII, etc.), presentation attributes (indentation, boldface type, etc.), and/or a uniform "template" of all elements (populated or not). The GILS record content analysis addresses this issue in more detail (see Appendix E-2 Record Content Analysis Findings, Discussion, and Recommendations).

S32g. [After 1 hour's experience comprising 6 scripted searches on GPO Access' and EPA's GILS plus up to 10 minutes nonscripted exploration of up to 5 nonbrokered GILS]

I will use GILS to locate government information in the future.

Findings: 9 of 10 participants answered this question.

Strongly Agree	Agree	Disagree	Strongly Disagree	No Opinion
1	5	3	0	0

Discussion: On the average, participants anticipate using GILS again. The "strongly agree" response was from the same participant "strongly agree"ing that "GILS is an efficient service" in S32h; in fact, all but 1 participant's responses mapped positively from future-use to efficiency of service.

Recommendations: User assessments should be followed up with a questionnaire concerning subsequent GILS vs. other government NIDR tool usage, discussion (positive or negative) of GILS with others, etc. within 6 to 8 weeks of the original session.

S32h. [After 1 hour's experience comprising 6 scripted searches on GPO Access' and EPA's GILS plus up to 10 minutes nonscripted exploration of up to 5 nonbrokered GILS]

GILS is an efficient service.

Findings: 9 of 10 participants answered this question.

Strongly Agree	Agree	Disagree	Strongly Disagree	No Opinion
2	3	4	0	0

Discussion: By a narrow call, participants considered GILS "efficient." It is noteworthy, however, that no one "strongly disagreed" to this question. The results of this question do not appear to correlate with those of S32l (GILS is an improvement over microfiche and paper indexes), indicating that participants judged "efficiency" relative to other networked information resources.

Recommendations: Further research could operationalize "efficient" for various user communities and types of information need; see S19.

S32i. [After 1 hour's experience comprising 6 scripted searches on GPO Access' and EPA's GILS plus up to 10 minutes nonscripted exploration of up to 5 nonbrokered GILS]

GILS probably helps agencies manage their information resources.

Findings: 9 of 10 participants answered this question.

Strongly Agree	Agree	Disagree	Strongly Disagree	No Opinion
1	4	1	0	3

Discussion: Bearing in mind that users did not encounter GILS policy documentation, it may be assumed that exposure to certain element names such as "control identifier" and "schedule number" led some participants to perceive an IRM objective within GILS.

Recommendations: None.

S32j. [After 1 hour's experience comprising 6 scripted searches on GPO Access' and EPA's GILS plus up to 10 minutes nonscripted exploration of up to 5 nonbrokered GILS]

It is clear to me how agencies choose what to include in GILS.

Findings: 9 of 10 participants answered this question.

Strongly Agree	Agree	Disagree	Strongly Disagree	No Opinion
0	0	2	3	4

Discussion: Participants were not clear as to the criteria for inclusion of an information resource in GILS collections. Of the Likert scale questions at the end of the script, this question evoked the strongest negative response. See also S26 and S30b,c,d,e.

Recommendations: See S26.

S32k. [After 1 hour's experience comprising 6 scripted searches on GPO Access' and EPA's GILS plus up to 10 minutes nonscripted exploration of up to 5 nonbrokered GILS]

Public electronic access to government information is important.

Findings: 9 of 10 participants answered this question.

Strongly Agree	Agree	Disagree	Strongly Disagree	No Opinion
8	1	0	0	0

Discussion: Our well-educated study demographic (P5), expressing a variety of government information needs (P6) clearly places importance on public access.

Recommendations: See S22; key to "access" is knowledge of availability. GILS marketing efforts should promote "the right to know."

S32l. [After 1 hour's experience comprising 6 scripted searches on GPO Access' and EPA's GILS plus up to 10 minutes nonscripted exploration of up to 5 nonbrokered GILS]

GILS is an improvement over microfiche and paper indexes.

Findings: 9 of 10 participants answered this question.

Strongly Agree	Agree	Disagree	Strongly Disagree	No Opinion
5	2	0	0	2

Discussion: Study participants consider GILS preferable to microfiche and paper indexes. The 2 users expressing "no opinion" searched both paper and online sources of government information relatively infrequently (P2a and P2b).

Recommendations: Further research could determine what features make GILS a more appealing locator tool—i.e., the ability to search across information providers (GPO Access GILS "All records" option), the ease of accessing GILS via Internet vs. a visit to a physical library, the degree of indexing, etc. These features should then be predominant in GILS marketing.

S32m. [After 1 hour's experience comprising 6 scripted searches on GPO Access' and EPA's GILS plus up to 10 minutes nonscripted exploration of up to 5 nonbrokered GILS]

All GILS records should contain information in all fields.

Findings: 9 of 10 participants answered this question.

Strongly Agree	Agree	Disagree	Strongly Disagree	No Opinion
3	2	3	0	1

Discussion: This result indicates a fairly strong consensus that all GILS elements should be populated, but no inferences can be drawn from the current study as to why. No correlation is apparent between participants' support of field-population and their self-confidence in fielded searching (S22).

Recommendations: It is recommended that GILS designers perform a confirmatory analysis that fielded searching improves retrieval precision in GILS for both known-item and exploratory searching.

S32n. [After 1 hour's experience comprising 6 scripted searches on GPO Access' and EPA's GILS plus up to 10 minutes nonscripted exploration of up to 5 nonbrokered GILS]

The quality of the records I examined varied widely.

Findings: 9 of 10 participants answered this question.

Strongly Agree	Agree	Disagree	Strongly Disagree	No Opinion
3	3	1	0	2

Discussion: One-third of respondents to this questions reported a wide variation in record quality. The GILS record content analysis addresses issues of quality in depth (see Appendix E-2 Record Content Analysis Findings, Discussion, and Recommendations).

Recommendations: Additional user-centered research is encouraged in order to operationalize "quality" criteria and develop and/or choose existing records as model(s) for benchmarking.

5.0. SUMMARY OF RECOMMENDATIONS

The following recommendations have been extracted from Section 4.0 Findings, Discussion, and Recommendations, and are arranged according to opportunities for increasing GILS acceptance among user communities. Many of the proposals can be implemented at the agency level, while others demand inter-agency consensus and cooperation. It is strongly recommended that regardless of the level of effort, agencies work toward adopting standard practices for GILS service features and record characteristics to enhance users' orientation in information space and promote GILS as a government-wide program.

5.1. Increase Users' Searching Confidence

- Present concise, comprehensible search instructions on the same page as the search input mechanism. Provide an example.
- Subject the element definitions presented to users to reading level (e.g., Flesch Reading Ease, Flesch-Kincaid Grade Level, etc.) and grammar checks to compute "fog index" by identification of incomplete clauses, jargon, passive voice, characters/syllables per word and words per sentence.
- Implement a record input mask or prevent-duplicates indexing feature to avoid application of the same Title or Control Identifier to more than one record.
- Avoid use of "complete text" and "full text" terminology in search options; use "all fields" instead.
- Provide a prominent link from agency GILS "index.htm" or search page to their mission statement and/or a list of general functions (perhaps from the *U.S. Government Manual*) to reduce dependency on an assumption of user cognizance.
- State clearly and prominently on each GILS site (ref: FedWorld) "Please also note that GILS records are intended to allow you to learn about what government information is available, not to be the information that you might be seeking!"
- Avoid linking to, or quoting verbatim, Federal information policy documentation for the purpose of introducing users to GILS functionality.

5.2. Improve GILS Niche In Information Space

- Entitle all record-source (nonbrokered) agencies' sites "[Agency/Bureau/Etc. Name] GILS" rather than "GILS" or in some manner make apparent that they are a component of the GILS universe.
- Include a straightforward statement on the GPO Access GILS search page(s) to the effect that the databases contain an agency's GILS records of information resources created and available from that agency and that GPO GILS listing may be incomplete.
- Determine the most effective way to convey GILS in the singular (as an agency database) vs. GILS as the collective.
- Provide a GILS hyperlink from popular government-information-seeking starting points (as determined by log transaction analysis, but we may assume the White House homepage, Library of Congress homepage, and GPO Access for examples).
- Create an information-space map of GILS that shows host overlaps and organizational relationships.
- State the scope of each GILS record collection in terms of, among other characteristics, organizational boundaries on each GILS site.
- Implement GILS on Z39.50-compliant servers and increase description of online resources to promote a perception of "seamless" service.

- Advertise in government publications and libraries; incorporate GILS linkage into agency and White House homepages; register/index GILS homepages with popular web search engines.
- Capitalize on GILS "exclusivity" from regular webpages (i.e., state that *only* through use of GILS can users directly access descriptions of thousands of agency resources).

5.3. Improve GILS Efficacy in NIDR and Revise NARA *Guidelines* Accordingly

- Develop and require a Resource Description element whose value is recognizable to the user, rather than the distributor, and is drawn from a controlled thesaurus; Appendix C-4 Record Content Analysis Methodology provides for one approach to operationalizing information object/container terminology .
- Adopt more precise nomenclature for Date of Last Modification (e.g., "GILS Record Revision Date" and "Time Period of Content" (e.g., "Time Span of Featured Subject Matter" vs. "Publication Date").
- Re-specify or abandon the concept of "Core" records, as it is the sole yardstick for grasping the extent of Federal information resource holdings described in GILS.

5.4. Improve the Quality and Consistency of GILS Records

- Enlist objective content reviewers to evaluate conformance of fields' content to qualitative descriptions and examples provided in the NARA *Guidelines*.
- Select or develop "model" records, as determined through user-satisfaction studies, to serve as an example to record creators and as a benchmark for evaluators.
- Enlist an objective party to evaluate hyperlinks to ensure that the record creator is not over-assuming an intellectual "hop" *vis-a-vis* the content of the linked-to page.
- View product prior to mounting as displayed by browser(s) selected on the basis of published usage reports (e.g., "Browser Battle" July 1996 *Internet World* p. 40) or agency website access log analysis.

6.0. AREAS FOR FURTHER RESEARCH

Results of the online scripted user assessment presented the following areas of exploration to increase GILS responsiveness and operational serviceability for users.

- Determine user satisfaction with agency name as the primary access point or as a starting point in exploring subject-oriented access to GILS systems
- Elucidate relationships between longevity of Internet experience, self-ratings of "familiarity" or "proficiency," and satisfaction with system results to inform the level of online help or hands-on training offered.
- Investigate the nature of users' information needs for government information by questions such "what was on your mind the last time you recall deliberately searching for or monitoring government information." For example, users seeking to satisfy a specific and direct but occasional information need may prefer the approach planned by USPS's WINGS (Web Interactive Network of Government Services) <[- Analyze the extent to which fielded searching in GILS improves for both known-item and exploratory searching.](http://www.wings.usps.gov>Topten>, which will present information on, among other things, tax-return filing, requesting birth certificates, and employment opportunities.• Perform server transaction log analysis to inform decisions about presentation of metadata elements for fielded searching and results.• Survey the agencies that are using controlled vocabulary and determine, through log transaction analysis, whether the practice increases retrievals.• Capture GILS user reactions, attributions of error/failure cause, and use of metadata elements through)

- Determine whether diversity in record display characteristics (file format, layout, etc.), results presentation, and searching features among agencies is a strength or weakness for end-users.
- Follow up all user sessions with a questionnaire concerning subsequent GILS vs. other government NIDR tool usage, discussion (positive or negative) of GILS with others, etc. within 6 to 8 weeks of the original session.
- Determine what features make GILS a more appealing locator tool (i.e., the ability to search across information providers [GPO Access GILS "All records" option]), the ease of accessing GILS via Internet vs. a visit to a physical library, the degree of indexing, etc.—and capitalize on these features in marketing.

7.0. CONCLUSION

The online scripted user assessments produced data that confirmed several important findings from other public user-oriented data collection activities such as focus groups.

GILS is not perceived as easy to use, predictable, or efficient in terms of satisfying information needs. Users perceive the use of bibliographic terminology and the lack of straightforward search instructions as uninviting, the content of the databases as unknowable, the service and record quality as uneven, and the lack of full text as approaching unforgivable.

Nonetheless, the results of the sessions and subsequent debriefings show that users believe in GILS potential if not all its current implementation characteristics. Further deployment of scripts to assess factors evoking both delight and disappointment among a variety of user communities—including librarians, publishers, public-information activists, business owners, and researchers—is recommended as an effective mechanism for gaining feedback from the “front line” in order to achieve that potential.

**Attachment E3-1
GILS Record from Script Item S28**

Title:

Index to the Wetland Educational Resources distributed by the New England regional office, EPA

Acronym:

Not supplied

Originator:

Environmental Protection Agency

Region 1

Wetland Protection Section

Controlled Vocabulary (Library of Congress Subject Headings):

Aquatic ecology; Conservation of natural resources; Biological diversity conservation; Biotic communities; Document delivery; Ecology; Environmental education; Environmental protection; Government publications; Wetlands; Wildlife

Controlled Vocabulary (Terms of Environment):

Habitat

Controlled Vocabulary (Supplied by GILS cataloger):

Educational materials

Local Subject Term:

US Federal GILS; wetlands; habitat protection; biodiversity

Abstract:

The U.S. Environmental Protection Agency's New England regional office produces and distributes numerous wetlands educational materials to the public — booklets, fact sheets, videocassettes, posters, etc. These materials cover a number of topics — what wetlands are, why they are important, how citizens (particularly students) can protect wetlands and the federal regulations that protect them. Materials are produced with specific target audiences in mind, including students, teachers, municipal officials, and developers.

Purpose:

The wetlands educational materials have been developed to increase the public's awareness of the importance of wetlands and how federal wetland regulations protect wetlands.

Agency Program:

Not supplied

Spatial Reference:

Geographic Keyword Name (Library of Congress Subject Headings):

New England

Spatial Reference:

Geographic Name (Hazardous Waste Superfund Database Thesaurus):

Region 1

Time Period of Content:

Time Period-Structured:

Not supplied

Time Period-Textual:

370

Not supplied

Availability:

Distributor:

Name: Wetland Protection Section

Organization: U.S. Environmental Protection Agency

New England Regional Office

Street Address: JFK Federal Building

City: Boston

State: MA

Zip Code: 02203

Country: USA

Hours of Service: 8:30 a.m. - 5:00 p.m. (EST) M - F

Telephone: 617-565-4421

Fax: 617-565-4940

Resource Description:

Not supplied

Order Process:

Materials may be ordered for free by writing to the address listed above.

Technical Prerequisites:

None

Available Linkage:

Not supplied

Available Linkage Type:

Not supplied

Sources of Data:

U.S. Environmental Protection Agency, environmental education nonprofit organizations, U.S. Army Corps of Engineers, other federal agencies

Access Constraints:

None

Use Constraints:

None

Point of Contact:

Name: Wetland Protection Section

Organization: U.S. Environmental Protection Agency

New England regional office

Street Address: JFK Federal Building

City: Boston

State: MA

Zip Code: 02203

Country: USA

Network Address: Not supplied

Hours of Service: 8:30 a.m. - 5:00 p.m. (EST) M - F

Telephone: 617-565-4421

Fax: 617-565-4940

371

Supplemental Information:

Information collection has a particular emphasis on wetlands stewardship materials for educators teaching grades K-12. All materials are located at the EPA New England regional office building, 1 Congress St., Boston, MA. For information on wetlands materials available throughout the country, contact the U.S. Environmental Protection Agency wetlands hotline at (800) 832-7828.

Schedule Number:

Not applicable

Control Identifier:

EPA/GENERAL01003

Record Source:

Environmental Protection Agency

Region 1

Wetland Protection Section

Date of Last Modification:

19951206

URL: <http://www.epa.gov/earth100/records/g01003.html>

**APPENDIX E-4
WEB SERVER TRANSACTION LOG ANALYSIS RESULTS**

APPENDIX E-4

WEB SERVER TRANSACTION LOG ANALYSIS RESULTS

1.0. INTRODUCTION

Use of the Internet in general, and the Web in particular, continues to increase dramatically. Indeed, as of January 1997, there are 16,146,000 Internet-based hosts and 828,000 domains (Network Wizards, 1997). This is nearly double the number of hosts and triple the number of domains as compared to January 1996 (Network Wizards, 1997).

Along with for-profit and non-profit organizations, Federal government agencies are increasing their use and provision of electronic networked services. Moreover, agencies continue to devote additional resources to the development and maintenance of Web-based services. Several critical Web service-related questions face the providers of such services:

- What is the server's traffic and overall ability and necessary resources to meet the demands of that traffic?
- What is the server's user community, as identified by the
 - accessing host IP address?
 - type of browser and operating system?
- What did users do while interacting with the server?
- From where did a user access and at what point leave the server?
- What problems did users encounter during their server sessions?

One means through which agencies can begin to answer these questions is through the analysis of Web-server generated log files. This appendix presents an overview of Web statistics, the process of analyzing and interpreting Web log files, and methodological details and findings of the research team's analysis of 14 days of Web server log files from the Environmental Protection Agency (EPA).

2.0. WEB SERVER LOG ANALYSIS: STEPS AND TOOLS

Log analysis is essentially a three step process that involves planning, data analysis, and interpretation activities. In particular, there is a need to:

- **Determine what types of information server administrators and decision makers need.** Log analysis is one means through which to determine whether Web-based services are meeting their intended missions or objectives. As such, server administrators and decision makers need to know what types of information are wanted prior to the analysis of Web server log files so as to ensure the collection of data that will assist in assessing mission or goal attainment.
- **Develop a program that can parse through, manipulate, and present value-added information from the log files.** Server administrators have the option of writing their own programs, downloading free software, or purchasing one of many off-the-shelf analysis products to do this. A listing of numerous Web analysis programs is available from webreference.com <<http://www.webreference.com/usage.html>>. Although continually increasing in their analysis capabilities, most programs tend to only parse through specific variables, leaving many important pieces of information untouched.
- **Analyze the information generated from the program.** This is not as straight-forward a process as one might think. For example, most log analysis software programs analyze the number of "hits" -- not accesses -- a server receives. In this case, the hit count reflects the number of items (e.g., images) downloaded when a user accesses a particular page. So, if a site has a corporate logo image file on every page, that image will more than likely be the most frequently downloaded -- "hit" -- item on the site. Analysis information such as that is relatively useless in determining the site's actual usage.

The program selection and analysis processes complement each other. Depending on the log analysis software used, server administrators are limited to certain types of log analysis. Based on those analysis limitations, server administrators need to know the meaning of the log analysis output (e.g., whether the statistics represent "hits" or "accesses.").

There are several Web server analysis software packages available on the market today, both free and for fees. Readers interested in reviewing sample Web analysis software should refer to the following sites:
<<http://www.statslab.cam.ac.uk/~sretl/analog/>>; <<http://www.mkstats.com>>;
<<http://www.ics.uci.edu/pub/websoft/wwwstat/>>; and, <<http://www.boutell.com/wusage/intro2.html>>. In addition, readers should review material found in Stout (1997). Readers may find the above listed log analysis software of use, depending on their analysis needs and requirements.

As discussed in the Methods section below, the study team assessed the currently available log analysis software packages and found them inadequate to perform the various analysis of the EPA log files of interest to the GILS evaluation study. The study team, therefore, developed its own PERL-based analysis scripts to analyze the EPA log files.

2.1. Background on Web Server Log Files

Web servers automatically generate and dynamically update four usage log files. These four log files and types of information each captures are as follows:

- Access Log (e.g., hits);
- Agent Log (e.g., browser, operating system);
- Error Log (e.g., download aborts); and
- Referer Log (e.g., referring links).

The log files are text files that can range in size from 1KB to 100MB, depending on the traffic at a given a web site (for additional background information, refer to Rubin, 1996; Noonan, 1996; Novak and Hoffman, 1996).

In determining the amount of traffic a site receives during a specified period of time, it is important to understand what, exactly, the log files are counting and tracking. In particular, there is a critical distinction between a *hit* and an *access*, wherein:

- A *hit* is any file from a web site that a user downloads. A hit can be a text document, image, movie, or a sound file. If a user downloads a web page that has 6 images on it, then that user "hit" the web site seven times (6 images + 1 text page).
- An *access*, or sometimes called a page hit, is an entire page downloaded by a user regardless of the number of images, sounds, or movies. If a user downloads a web page that has 6 images on it, then that user just accessed one page of the web site.

This distinction is noteworthy. Most web analysis software counts the number of *hits* a server receives, rather than the number of *accesses*.

2.1.1. Access Log

The Access Log provides the greatest amount of server data, including the date, time, IP address, and user action (e.g., document/image/sound/movie download). The following is an example line of text from an Access Log:

smx-ca8-50.ix.netcom.com - - [30/Sep/1996:02:57:07 -0400] "GET/Proj/main.html"

It is possible to analyze the following variables in the Access Log:

- **Domain name or Internet Protocol (IP) number.** In the above example we know that the user's computer had the following domain name: smx-ca8-50.ix.netcom.com.
- **Date and Time.** In the above example we know that the user accessed a page on September 30, 1996 at 2:57 AM and 07 seconds. By default the time is based on a twenty four hour clock.

- **Item accessed.** The word item can mean an image, movie, sound, or html file. The above example shows that main.html was the item accessed. It is also important to note that the full path name (from document root) is given to avoid confusion, (e.g., there may be more than one main.html on a server).

It is possible to generate the following data from these variables,::

- The percentage of users accessing the site from a specific domain type (e.g., .com, .edu, .net, .mil, .gov). This can be analyzed further by *hits* versus *accesses*.
- The number of hits the server is getting from various IP groups. Such data can inform server administrators as to the primary clients of their servers.
- The number of unique IP addresses accessing the site. While not a measure of unique users, this can provide server administrators with some indication of the number of users by stripping repeat IP addresses from the log data. This data is an important indicator of the breadth of penetration of a server.
- The number of accesses/hits the server receives during specific hours and days of the week. These statistics can be useful to server administrators who need to know the optimal time/day to perform server maintenance and/or upgrades.
- The path -- known as "threading" -- a user takes through a site. Knowing this allows server administrators to determine the average length of a user's session, specific location duration (e.g., average time on a page), average download times, and how the user navigated through the site (e.g., entrance and exit points).

The data from the Access Log provides a broad view of a Web server's use and users (as indicated by IP addresses). Such analysis enables server administrators and decision makers to characterize their server's audience and usage patterns.

2.1.2. Agent Log

The Agent Log provides data on a user's browser, browser version, and operating system. This is significant information, as the type of browser and operating system determine what a user is able to access on a site (e.g., Java, forms). Below is a sample Agent Log entry:

Mozilla/3.0 (Win95; I)

Analysis of the Agent Log enables server administrators to determine the (see Figures 6-8):

- **Browser.** The type of browser used to access a web site. There are several different Web browsers on the market today (e.g., Netscape, Microsoft Internet Explorer, LYNX, Mosaic), each of which have different viewing capabilities.
- **Browser version.** The version of a browser used. Not all browsers can view all components of a Web site. For example, Netscape version 1.0 cannot view forms-based data.
- **Operating system.** The type of computer and operating system users have. A Web site can look different to users depending on their computer platform (e.g., Windows, Win95, Macintosh, PowerPC, SunOS).

These data are essential for the design and development of Web sites. Without such information, server administrators could design sites that require viewing capabilities that a vast majority of the site's users do not possess. At best, this leads to wasted effort by the server administrator. At worst, this can lead to improperly displayed Web content, thus effectively rendering the site useless to the user.

2.1.3. Error Log

The average Web user will receive an "Error 404 File Not Found" message several times a day. When a user encounters this message, an entry is made in the Error Log. Below is a sample Error Log entry:

[Sun Nov 3 23:57:00 1996] httpd: send aborted for pm02_23.ct.net,
URL:/OWOW/images/new/owpool.gif

The Error Log contains the following data for analysis:

- **Error 404.** The Error Log tells a server administrator the time, domain name of the user, and page on which a user received the error. These error messages are critical to Web server administration activities, as they inform server administrators of problematic and erroneous links on their servers.
- **Stopped transmission.** This informs a server administrator of a user-interrupted transfer. For example, a user clicking on the "stop" button would generate a "stopped transmission" error message. The Error Log tells a server administrator the time, domain name, and page that a user was on when the transmission was stopped (as in the above sample Error Log entry). This information is useful as it can indicate patterns with large files such as image, movie, and other files that users consistently stop downloading.

The analysis of Error Log data can provide important server information such as missing files, erroneous links, and aborted downloads. This information can enable server administrators to modify and correct server content, thus decreasing the number of errors users encounter while navigating a site.

2.1.4. Referer Log

The Referer Log indicates what other sites on the Web link to a particular server. Each link made to a site generates a Referal Log entry, a sample of which is below:

http://www.altavista.digital.com/cgi-bin/query?pg=q&what=web&fmt=.&q=SIC+CODE->/xxx/html/rcri/rcri_sic_code.html

In this particular example, the referer was AltaVista, indicating that the user entered the Web site after performing a search using the AltaVista search facility.

The Referer Log entry provides the following data:

- **Referral.** If a user is on a site (e.g., ericir.syr.edu), and clicks on a link to another site (e.g., www.sun.com), then www.sun.com will receive an entry in their Referer Log. The log will show that the user came to the sun site (www.sun.com) via ericir.syr.edu (the referral).

Such referral data is critical to alleviating missing link (Error 404) data. For example, when the URL of a page within www.sun.com changes, the server administrator of www.sun.com could notify all referrals (e.g., ericir.syr.edu) of the change. This can alleviate future "Error 404 - File Not Found" messages.

Through the analysis of the four log files, Web service providers can begin the process of assessing and evaluating their networked information services. Current Web usage statistics generally center on the analysis of the Access Log, thus limiting the ability of Web-mounted service extensiveness measures. There are, however, means to analyze the Agent, Error, and Referer log files. Such techniques can provide important additional insight into the use of Web-based services by users.

3.0. METHODOLOGY

As part of the evaluation study of U.S. Implementation of GILS, the authors selected one Federal agency's Web server from which to collect log files. The authors performed analysis on a sample of the log files to:

- Determine the overall Web site's traffic, including the
 - origin of users
 - portions of the site that are accessed
 - number of document downloads (both hits and accesses);
- Determine the use of the Web site GILS traffic, including the
 - origin of users

- portions of the site that are accessed
- number of document downloads (both hits and accesses);
- Experiment with developing new log analysis techniques that go beyond domain, hit, and browser counts;
- Assist Federal agencies that operate Web-based GILS servers to develop, implement, and maintain on-going log file analysis; and
- Inform Federal agencies that operate Web-based GILS servers of the utility in analyzing and interpreting log file data in on-going assessments of their GILS implementations.

Such an evaluation enables the maintainers, policy makers, and stakeholders of the Web site to determine a site's use as one component of an overall networked information resource.

The log files were collected daily between February 2, 1997 and February 15, 1997. The four log files ranged in size from 8 megabytes to 26 megabytes each per day. In all, approximately 560 megabytes of log file data were collected. The resulting output, Web log file analysis PERL scripts, and log files together consumed approximately 1 gigabit of storage. The analysis of the EPA log files was performed on a Pentium 150 MHZ computer with 32 MB of RAM, and the analysis of each of the four daily log files took approximately 40 minutes.

3.1. Choosing Web Analysis Software

The authors reviewed multiple Web analysis software packages along the following criteria:

- Ability to provide global and directory specific Web server analysis;
- Ability to distinguish between hits and accesses;
- Ability to determine user-specific actions (e.g., navigation) through a Web site session; and
- Ability to distinguish between unique and total referrals.

Most existing log analysis software could perform one or more of the above functions. None, however, met all the analysis criteria for the GILS evaluation project. As such, the authors worked with a study team at the School of Information Studies, Syracuse University, to develop PERL-based Web analysis scripts that would provide all the required analysis capabilities. Readers, therefore, will find that currently available Web log file analysis software cannot perform some of the analysis techniques presented in this appendix.

3.2. Developing the PERL Scripts

The development and pre-testing of the PERL scripts required considerable effort. The Syracuse University script development team required the equivalent of 240 man-hours developing the scripts. An additional 100 man-hours were required to pre-test the scripts using several different log files from different servers, including a test data set from the Federal agency HTTP GILS server. Running the scripts on the 14 day period of EPA log files and outputting the analysis into a usable format required an additional 100 man-hours. In total, therefore, the PERL script development process consumed approximately 420 man-hours.

To ensure valid and reliable results, script file results were compared to results generated by other log analysis software, where possible. When errors in script files were found, corrections were made and the files re-tested.

4.0. FINDINGS

The study team analyzed each of the four log files on a daily basis. Analysis of the files is presented both in aggregate and individual-day format where possible. To simplify the presentation of the data, the findings are presented by log file type. Readers should note that these findings do not include such commonly available analysis as hits by time of day or day or week. Rather, this appendix presents findings from the use of the developed PERL scripts intended to provide new and previously unavailable forms of log analysis.

4.1. Access Log

The EPA Web server generates considerable traffic on any given day (see Tables 1 and 2). On average, the EPA server receives approximately 80,000 daily accesses that generate over 213,000 daily hits (see Tables 1 and 2). In all, the EPA server received over 564,000 accesses per week generating over 1,496,000 hits per week. As Tables 1 and 2 demonstrate, the EPA server is most used during the middle part of the week.

On average, the GILS component of the EPA server (As measured by use of the Earth100 directory), the daily percentage of GILS accesses ranges from .45% to .93%, with a weekly average of .52% and .61%, respectively (see Tables 3 and 4). These GILS accesses account for .20% to .44% of all EPA server hits. As with the EPA server in general, the GILS portion of the server is most heavily used during the middle of the week.

It is important to note three factors when considering the average GILS usage patterns as depicted in Tables 3 and 4:

- The tables do not include Z39.50 accesses to GILS records, thus do not necessarily reflect the total usage of the EPA GILS database; and
- The tables do not compare, nor did the study collect such data, the traffic the GILS component of the EPA server to other EPA server components. The EPA server has a significant number of subdirectories that would require traffic analysis to gain a more accurate sense of the GILS directory traffic in relation to other server directories.

The overall use of the EPA GILS records is underreported in Tables 3 and 4 without such data.

It is interesting to note that the EPA server in general, and the GILS portion in particular, both receive a fairly consistent percentage of traffic (as measured by accesses and hits) from within the United States and from foreign countries (see Tables 5 and 6). The daily average for United States-generated EPA accesses ranges from 72.48% to 74.19%, while the daily average accesses from foreign countries ranges from 25.81% to 27.52% (see Tables 5 and 6). The daily average for United States-generated GILS accesses ranges from 28.39% to 35.06%, while the daily average accesses from foreign countries ranges from 28.39% to 35.06% (see Tables 5 and 6). From the limited data set, it is not possible to state whether foreign country access to EPA GILS data is on the rise, as the rise in Table 6 demonstrates.

Perhaps one of the more innovative log analysis techniques developed for this study is that of path analysis. Path analysis enables a Web server administrator to determine a user's path and actions through a server for any given session. Table 7 demonstrates the possibility of such an analysis technique (the full IP address of the user was removed to protect the identity of that user). As Table 7 demonstrates, user xxx.oln.com first accessed the EPA server at the server's home page at 7:44AM. The user accessed a variety of EPA pages and generated several hits. At 7:46AM, the user entered GILS Earth100 directory, accessed a variety of files, and then exited the Earth100 directory at 7:54AM. The user remained logged onto the server, but did not perform any additional actions until 12:38PM. At that time, the user browsed and performed a variety of searches before logging off the server at 12:43PM. In all, the user remained logged onto the EPA server for 5 hours and 59 minutes.

In summary, it is possible to demonstrate the following through the EPA access log data:

- The total number of daily and weekly accesses and hits;
- The average daily accesses and hits;
- The percentage of accesses and hits generated by EPA's GILS, as measured through the Earth100 directory;
- The percentage of accesses and hits that the EPA server in general and EPA's GILS in particular (as measured through the Earth100 directory) experienced from the US and foreign countries; and
- The specific path a user takes through a server per session.

Together, these statistics provide an understanding of the overall use of the EPA Web server, as well as particular information resources (e.g., GILS) provided by the server.

Table 1. EPA Server Number of Hits/Accesses for February 2, 1997 to February 8, 1997.

Date	Hits	Accesses
2-Feb	131,094	42,461
3-Feb	296,687	100,437
4-Feb	307,714	104,321
5-Feb	301,650	101,530
6-Feb	278,000	93,348
7-Feb	260,330	85,789
8-Feb	113,121	41,440
Week one totals	1,688,596	569,326
Week one averages	241,228	81,332

Table 2. EPA Server Number of Hits/Accesses for February 9, 1997 to February 15, 1997.

Date	Hits	Accesses
9-Feb	131,015	52,055
10-Feb	78,174	26,358
11-Feb	267,864	96,631
12-Feb	322,080	123,782
13-Feb	306,930	116,504
14-Feb	280,737	109,583
15-Feb	109,327	39,863
Week two totals	1,496,127	564,776
Week two averages	213,732	80,682

Table 3. EPA Server Earth100 Directory Number of Hits/Accesses for February 2, 1997 to February 8, 1997.

Date	Hits	% of EPA Server	Accesses	% of EPA Server
2-Feb	292	.22%	232	.55%
3-Feb	581	.20%	449	.45%
4-Feb	684	.22%	550	.53%
5-Feb	639	.21%	502	.49%
6-Feb	642	.23%	498	.53%
7-Feb	589	.23%	407	.47%
8-Feb	417	.37%	339	.82%
Week one totals	3,844	.23%	2,977	.52%
Week one averages	549	.23%	425	.52%

Table 4. EPA Server Earth100 Directory Number of Hits/Accesses for February 9, 1997 to February 15, 1997.

Date	Hits	% of EPA Server	Accesses	% of EPA Server
9-Feb	462	.35%	331	.64%
10-Feb	225	.29%	163	.62%
11-Feb	902	.34%	625	.65%
12-Feb	1,097	.34%	771	.62%
13-Feb	853	.28%	589	.51%
14-Feb	792	.28%	603	.55%
15-Feb	493	.44%	369	.93%
Week two totals	4,824	.32%	3,451	.61%
Week two averages	689	.32%	493	.61%

Table 5. EPA Server and Earth100 Directory Percentage of Accesses From Country of Origin for February 2, 1997 to February 8, 1997.

Date	EPA		GILS (Earth100)	
	US Accesses	Outside US	US Accesses	Outside US
2-Feb	73.09%	26.91%	76.34%	23.66%
3-Feb	71.07%	28.93%	62.96%	37.04%
4-Feb	70.98%	29.02%	68.25%	31.75%
5-Feb	70.96%	29.04%	70.70%	29.30%
6-Feb	71.63%	28.37%	70.86%	29.14%
7-Feb	71.50%	28.50%	69.91%	30.09%
8-Feb	78.11%	21.89%	82.22%	17.78%
Week one averages	72.48%	27.52%	71.61%	28.39%

Table 6. EPA Server and Earth100 Directory Percentage of Accesses From Country of Origin for February 9, 1997 to February 15, 1997.

Date	EPA		GILS (Earth100)	
	US Accesses	Outside US	US Accesses	Outside US
9-Feb	79.51%	20.49%	80.05%	19.95%
10-Feb	64.51%	35.49%	46.20%	53.80%
11-Feb	73.97%	26.03%	70.59%	29.41%
12-Feb	77.24%	22.76%	74.27%	25.73%
13-Feb	76.75%	23.25%	73.46%	26.54%
14-Feb	76.14%	23.86%	61.96%	38.04%
15-Feb	71.21%	28.79%	48.04%	51.96%
Week two averages	74.19%	25.81%	64.94%	35.06%

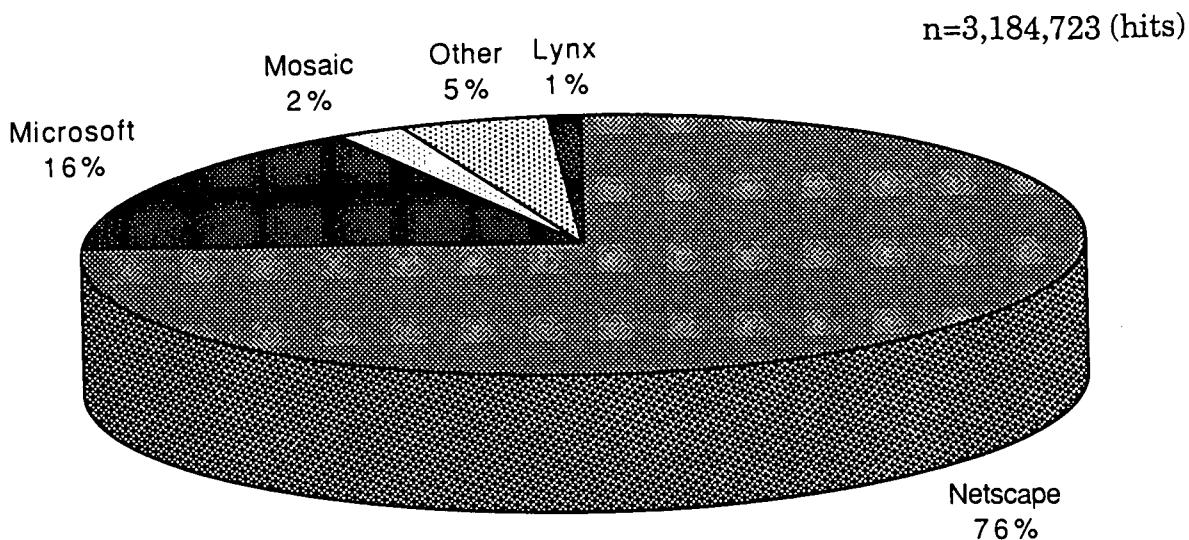
Table 7. EPA GILS Path Analysis for User xxx.oln.com for February 12, 1997.

Path and Time
/ 07:44:17
/epahome/images/2title1n.gif 07:44:21
/epahome/images/browse.gif 07:44:21
/epahome/images/newmenu.gif 07:44:21
/epahome/images/2message.gif 07:44:22
/epahome/images/newmenu.map?436,182 07:44:56
/epahome/finding.html 07:44:58
/epahome/images/epahr1.gif 07:45:04
/epahome/images/2searchn.gif 07:45:04
/Access/index.html 07:45:38
/Access/images/epaseal.gif 07:45:41
/icons/construction.gif 07:45:42
/earth100/ 07:46:43
/cgi-bin/odometer.gif?/gils.html&width=6&.gif 07:46:47
/earth100/fish.gif 07:46:48
/oar/images/exit.gif 07:46:48
/earth100/browse.html 07:47:17
/earth100/browse/index.html 07:47:28
/icons/epabar.gif 07:47:31
/earth100/browse/C.html 07:47:52
/earth100/records/a00108.html 07:48:29
/earth100/records/a00195.html 07:51:06
/earth100/browse/L.html 07:52:52
/earth100/browse/E.html 07:53:56
/earth100/browse/I.html 07:54:19
/ 12:38:47
/epahome/images/2message.gif 12:38:57
/epahome/images/newmenu.gif 12:38:57
/epahome/images/browse.gif 12:38:57
/epahome/images/2title1n.gif 12:38:57
/epahome/images/newmenu.map?310,37 12:39:27
/epahome/images/browse.map?391,16 12:39:46
/epahome/search.html 12:39:55
/epahome/search.html 12:39:56
/epahome/images/2title1n.gif 12:39:58

Table 7. EPA GILS Path Analysis for User xxx.oln.com for February 12, 1997.

Path and Time
/oar/images/exit.gif 12:39:58
/epahome/images/2searc1n.gif 12:39:58
/epahome/images/epahr1.gif 12:39:59
/cgi-bin/waisgateII 12:40:15
/epahome/images/epa.gif 12:40:25
/cgi-bin/waisgateII?WAISdocID=6921415498+30+0+0&WAISaction=retrieve 12:41:25
/epahome/mapping.htm 12:41:35
/epahome/images/2searc1n.gif 12:41:45
/epahome/images/epahr1.gif 12:41:45
/cgi-bin/waisgateII?WAISdocID=6921415498+30+0+0&WAISaction=retrieve 12:42:07
/epahome/images/2title1n.gif 12:42:15
/epahome/images/2searc1n.gif 12:42:18
/epahome/images/epahr1.gif 12:42:19
/cgi-bin/waisgateII?WAISdocID=6921415498+0+0+0&WAISaction=retrieve 12:42:32
/cgi-bin/waisgateII?WAISdocID=6921415498+0+0+0&WAISaction=retrieve 12:42:51
/icons/epabar.gif 12:42:58
/epahome/images/2title1n.gif 12:43:23
Arrived: 07:44:17 Left: 12:43:23
Total time: 5:59:05

**Figure 1. EPA Agent Log for February 2, 1997 to Febuary 15, 1997
by Type of Browser.**



4.2. Agent Log

From the agent log, it is possible to determine the type of browser that users use when accessing a particular site. As Figure 1 demonstrates, Netscape is the browser of choice for a vast majority of EPA server users with 76%, followed by Microsoft Internet Explorer with 16%, Other (e.g., AOL browser, GNN)¹ with 5%, Mosaic with 2%, and Lynx with 1%.

As Figure 2 indicates, the agent log can be further analyzed to determine what version of a particular browser used to access a site. A vast majority of users -- 70.49% -- use a version of Netscape that is 2.0 or later. Thus, most all users can access forms-based Web data (a feature incorporated into later versions of Netscape). Only 41.06% of Netscape users, however, can access Java-based Web data (a feature incorporated into Netscape 3.0).

Figure 3 shows that a vast majority of users accessing the EPA Web site use a PC platform with 47% using Windows, 35% using Windows95, and 2% using Windows NT. Mac users account for only 9% of the EPA server traffic.

In summary, the agent log data provides the following data:

- The type of browser used to access a Web site;
- The version of a browser used to access a Web site; and
- The operating system of the computer used to access a Web site.

These data are particularly important as they indicate to the Web server administrators the access and display capabilities of users.

¹To recognize all the various browsers, the PERL scripts need to specifically look for each browser. The scripts used for this analysis counted the major existing browsers.

Figure 2. EPA Agent Log for February 2, 1997 to Febuary 15, 1997 by Version of Netscape Browser.

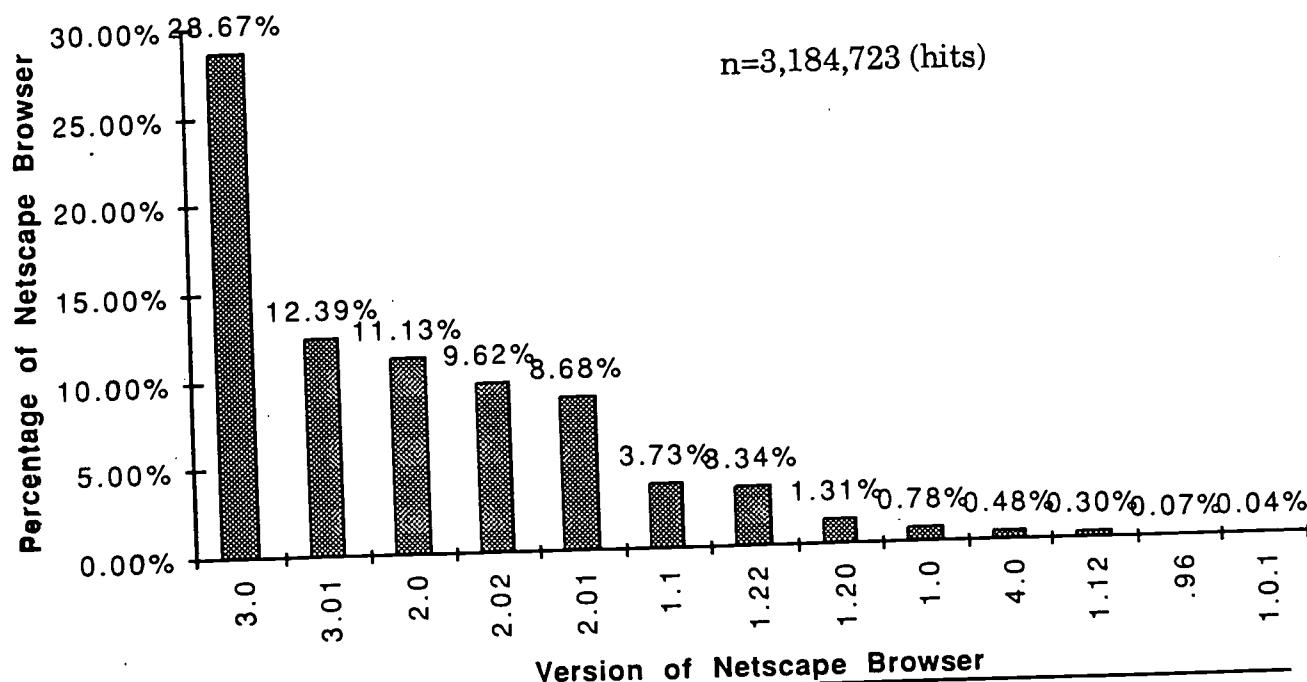
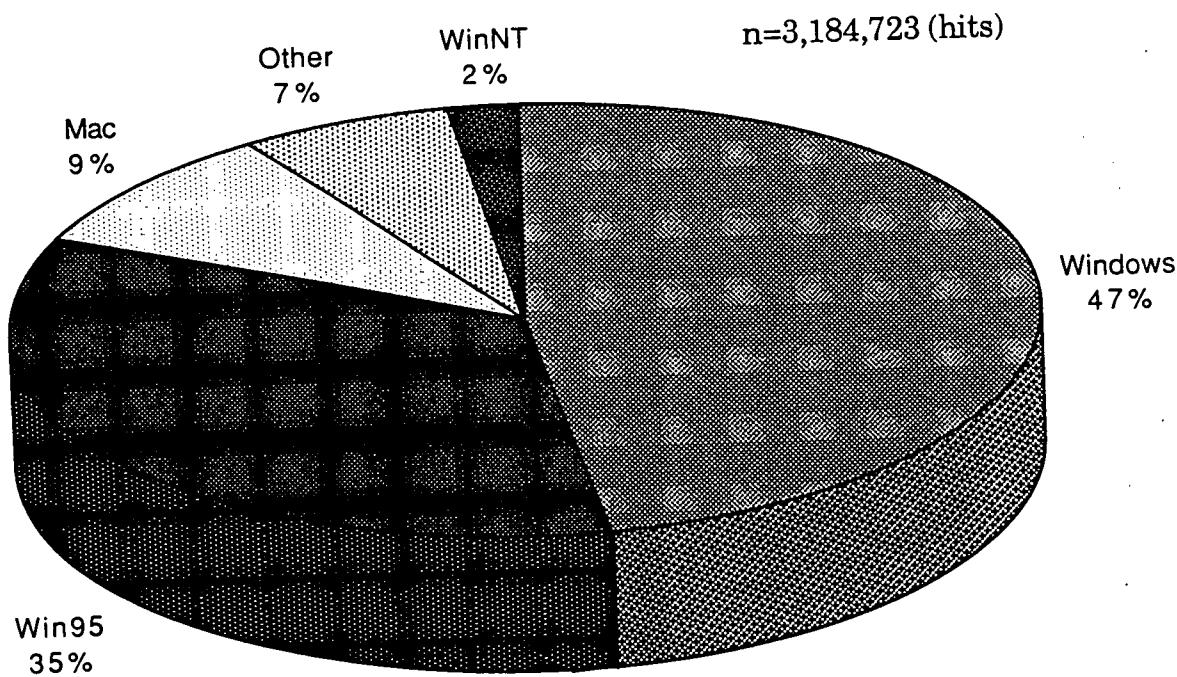


Figure 3. EPA Agent Log for February 2, 1997 to Febuary 15, 1997 by Type of Operating System.



4.3. Referer Log

The referer log indicates the Web site from which a user enters the Web server of interest. For example, if a user conducts a search using the AltaVista search engine and finds a retrieved search item of interest, the receiving Web site will show that AltaVista "referred" the user to their site. The referer log will also indicate the number of erroneous referrals (e.g., problematic links).

Tables 8 and 9 show that the EPA server received an average of 7,900 daily unique referrals (i.e., unique IP addresses) and an average of 11,500 total referrals (including multiple referrals from the same IP addresses). As indicated in Tables 8 and 9, the week of February 9, 1997 to February 15, 1997 had considerably more referral errors -- 29,937 as compared to 19,967 for the week of February 2, 1997 through February 8, 1997. It is not possible to determine precisely the reason for this increase in errors. Tables 8 and 9 also indicate that the GILS directory receives relatively few referrals as compared to the EPA server in general (a daily average of 5 referrals). It is interesting to note, however, that most of the GILS referrals are from unique IP addresses, indicating that users access the GILS directory from different sources each time.

The EPA server receives so many referrals that it is not feasible to identify all referring sources. Based on the analysis of the referer log, however, it was possible to identify the four most frequently referring sites to the EPA server (see Tables 10 and 11). The most frequently referring sites to the EPA server are the Yahoo-Society site with an average daily referral rate of 68 and 77 respectively, the Yahoo-Government site with an average daily referral rate of 60 and 66 respectively, the Web Directory with an average daily referral rate of 46 and 55 respectively, and the Web Crawler with an average daily referral rate of 30 and 39 respectively.

In summary, it is possible to generate the following data from the referer log:

- The total and unique number of referring IP addresses; and
- The most frequently referring sites.

This type of data enables Web server administrators to determine who their most frequent referring sites are as well as the overall number of referrals. Such data is particularly useful for Web server administrators as changes to document links are made on a server. With the assistance of the referer log data, server administrators of a site can contact the most frequently referring sites to make changes in their links so as to avoid user-encountered error messages.

Table 8. EPA Server and Earth100 Directory Referrals for February 2, 1997 to February 8, 1997.

Date	EPA			GILS (Earth100)	
	Total Referrals	Unique Referrals	Total Errors	Total Referrals	Unique Referrals
2-Feb	7,318	5,462	890	2	2
3-Feb	13,946	9,076	1,638	3	2
4-Feb	15,119	9,786	2,440	8	8
5-Feb	14,304	9,199	4,708	3	3
6-Feb	13,506	8,889	3,931	24	8
7-Feb	12,320	8,330	3,980	27	8
8-Feb	6,763	5,151	2,380	1	1
Week one totals	83,276	55,893	19,967	68	32
Week one averages	11,897	7,985	2,852	10	5

Table 9. EPA Server and Earth100 Directory Referrals for February 9, 1997 to February 15, 1997.

Date	EPA			GILS (Earth100)	
	Total Referrals	Unique Referrals	Total Errors	Total Referrals	Unique Referrals
9-Feb	7,570	5,644	2,657	2	2
10-Feb	13,777	9,058	4,396	7	7
11-Feb	13,978	9,370	5,045	9	8
12-Feb	13,941	9,484	5,089	3	3
13-Feb	14,022	9,262	5,707	7	7
14-Feb	11,647	7,885	4,513	5	5
15-Feb	6,061	4,628	2,530	5	5
Week two totals	80,996	55,331	29,937	38	37
Week two averages	11,571	7,904	4,277	5	5

Table 10. EPA Server Top Referring Sites for February 2, 1997 to February 8, 1997.

Date	Web Directory	Yahoo -Society	Yahoo-Government	Web Crawler
2-Feb	46	43	36	29
3-Feb	58	76	82	55
4-Feb	65	91	83	36
5-Feb	59	82	92	41
6-Feb	80	80	84	45
7-Feb	52	61	65	41
8-Feb	27	44	22	25
Week one totals	387	477	464	272
Week one averages	55	68	66	39

Table 11. EPA Server Top Referring Sites for February 9, 1997 to February 15, 1997.

Date	Web Directory	Yahoo -Society	Yahoo-Government	Web Crawler
9-Feb	36	60	37	28
10-Feb	75	89	91	46
11-Feb	53	82	79	28
12-Feb	49	89	68	24
13-Feb	33	101	70	35
14-Feb	42	67	54	32
15-Feb	33	48	20	19
Week two totals	321	536	419	212
Week two averages	46	77	60	30

4.4. Error Log

The error log provides data on the errors (e.g., dead links, aborted file downloads) that users either encounter or initiate. Due to the large volume of traffic that the EPA Web server generates, it is not feasible to present all the error log data. As such, selected error data is presented in this section.

Tables 12 and 13 show that six files are consistently aborted by users. These include the cgi-bin/waisgateII (an average daily download abort rate of approximately 200), cgi-bin/waisgate (an average daily download abort rate of approximately 160), /oar/oarmap.gif (an average daily download abort rate of approximately 50), /icons/nceri2.gif (an average daily download abort rate of approximately 40), epahome/404.html (an average daily download abort rate of approximately 35), and OW/images/feb_ad.gif (an average daily download abort rate of approximately 32).

Tables 14 and 15 demonstrate both the total number of download aborts as well as the most consistently aborted downloaded file for the GILS Earth100 directory. The data indicate that the GILS directory has a daily average of 5 send aborts, most of which are due to the isi.zip file (see Tables 14 and 15).

In summary, the error log data provide the following data:

- The overall and directory specific number of aborted downloads; and
- The overall and directory specific most frequently aborted file downloads.

This type of data is helpful to Web administrators as it indicates which images, files, and pages require too much time to download, leading users to abort the files.

The next section presents key issues that the study team encountered in developing the PERL scripts for EPA log analysis, the formatting of the EPA log files, and the general management of EPA's log files.

Table 12. EPA Server Most Frequently Aborted Downloads for February 2, 1997 to February 8, 1997.

Date	cgi-bin/ waisgate	cgi-bin/ waisgateII	/oar/ oarmap.gif	epahome /404.html	OW/images /feb_ad.gif	/icons /nceri2.gif
2-Feb	65	42	30	34	10	16
3-Feb	248	316	82	58	44	66
4-Feb	291	314	59	60	39	78
5-Feb	266	258	88	44	42	52
6-Feb	250	216	76	33	48	49
7-Feb	179	173	44	40	44	43
8-Feb	54	92	22	23	17	13
Week one totals	1,353	1,411	401	292	244	317
Week one averages	193	202	57	42	35	45

Table 13. EPA Server Most Frequently Aborted Downloads for February 9, 1997 to February 15, 1997.

Date	cgi-bin/ waisgate	cgi-bin/ waisgateII	/oar/ oarmap.gif	epahome /404.html	OW/images /feb_ad.gif	/icons /nceri2.gif
9-Feb	43	110	24	13	16	9
10-Feb	237	264	50	28	31	38
11-Feb	245	310	74	46	39	54
12-Feb	225	257	64	48	59	51
13-Feb	170	234	58	42	33	37
14-Feb	180	207	55	61	38	68
15-Feb	46	102	18	35	18	10
Week two totals	1,100	1,382	325	238	216	257
Week two averages	157	197	46	34	31	37

Table 14. EPA Server Total and Most Frequently Aborted Downloads for the Earth100 Directory for February 2, 1997 to February 8, 1997.

Date	Total send aborts	Send aborts on isi.zip
2-Feb	7	6
3-Feb	6	2
4-Feb	2	1
5-Feb	5	2
6-Feb	3	3
7-Feb	9	4
8-Feb	0	0
Week one totals	32	18
Week one averages	4.57	2.57

Table 15. EPA Server Total and Most Frequently Aborted Downloads for the Earth100 Directory for February 9, 1997 to February 15, 1997.

Date	Total send aborts	Send aborts on isi.zip
9-Feb	4	2
10-Feb	13	3
11-Feb	7	4
12-Feb	2	0
13-Feb	4	1
14-Feb	6	1
15-Feb	5	4
Week two totals	41	15
Week two averages	5.86	2.14

5.0. KEY DATA AND LOG FILE ANALYSIS ISSUES AND RECOMMENDATIONS

Throughout the log analysis process, the study team encountered a number of problems and issues that affected its ability to develop log analysis script files and perform certain types of log analysis. Below is a list of key issues encountered and recommendations to resolve such problems:

- *Transfer of files.* The study team had no guarantee that the files it received were the complete data set. There is a need to implement a procedure whereby EPA would post the file size of the log files directly from the server and the study team could verify this against the downloading files. An example of this is for the February 10, 1997, for which there was clearly data missing from the access log.
- *Storage space.* Storing just two weeks of log files from the EPA, as well as the PERL scripts and the resulting files took up nearly a gigabit of hard drive space. If the development and analysis of EPA's log files were to continue, the study team would need to dedicate a machine with adequate hard drive space to maintain the files. Moreover, as the study team suffered a server crash during this study, it is also necessary to have a back up server or tape backup of the script and log data files.
- *Enhancing the access log scripts.* Given time and resource constraints, a majority of the analysis for this project was done using Microsoft Excel. It is possible to incorporate some of this analysis into a re-write of the PERL scripts. For example, the percentages of U.S. hits versus outside hits in the access log were added manually. This calculation would be relatively simple to have the PERL scripts create.
- *Accommodating different log file formats.* A portion of the PERL scripts used for this project were originally written and tested with Syracuse University-generated log files. The study team found, however, that the EPA's Web server used substantially different file formats. For example, the original PERL scripts did not count files with .txt extensions or cgi-bin files as accesses, a feature required for the accurate analysis of the EPA server files. The study team re-wrote the scripts prior to the final analysis presented in this study. In the future, though, it would be necessary to have a team of people responsible for assessing the log file format and composition of the Web server prior to using the scripts to ensure that no file formats and/or other data were missing.
- *Awareness of script and counting errors.* The study team encountered two main errors in their PERL script development: (1) double counting hits, and (2) erroneous counting of EPA main page accesses. The study team initially developed PERL scripts to count the number of server hits as well as the percentage of accesses directly to the EPA server's main page. However, the scripts erroneously double counted some hits and could not accurately track the main page accesses. The hit counting errors were corrected, but the percentage of main page access could not be. The latter data were not reported in this study.
- *Separating in-house from external referrals.* There are a large number of EPA-based hosts that access the EPA server (e.g., www.epa.gov, ftp.epa.gov, earth1.epa.gov). It is, therefore, difficult to completely remove EPA-based accesses and hits from the referral logs. While the study team attempted to remove such internal referrals to get a more accurate picture of the server's non-EPA use, the study team is certain that they were unable to adjust for all the EPA domain names that might exist.
- *Separating various search engines.* The study team underestimated the number of daily referrals that the EPA server received. From analyzing the log analysis data, it is clear that a number of the referrals come from search engines. The PERL scripts were not written to extract this information. Future development of the scripts can help the study to determine not only what percentage of referrals come from search engines, but what search engines users tend to use and what search terms users enter.
- *Specific path and error analysis.* At present, the PERL scripts can only analyze a specific directory from the access log (e.g, the GILS [Earth100] directory). Incorporating error log information into the referer log analysis, however, required a custom shell script. More work is needed to fully incorporate error and referer log file data.
- *Extended log format.* There are multiple types of log file formats. The EPA Web server currently generates log files in the common log format. The extended log file format, however, allows all log information to be collected into one log file. Although this would mean the study team's PERL scripts

would require a complete revision, it would be possible to collect more information about specific visitors through this file format.

The above issues provide an insight into the key problems that the study team encountered and attempted to resolve while performing the log analysis of the EPA Web server log files. The problems illustrate the newness of log file analysis, the lack of consistency of log file formats, and the need to develop additional means of analyzing Web server log files.

The next section presents key issues in the collection, use, and interpretation of log file data.

6.0. KEY LOG FILE INTERPRETATION AND MANAGEMENT ISSUES

There are several key issues that Web service providers should consider when using log files as indicators of digital service output measures. These include:

- *Interpreting and considering the log files as one component of a larger assessment activity for networked services.* While log files can provide Web administrators and others with critical server-related data, log files do not reflect user-based impact and outcome measures. Log files, therefore, combine both user and technical perspectives on Web services.
- *Understanding what, exactly, the data reflect.* The distinction between "hits" (downloads on an html page) and accesses (a downloaded html page) is critical. Software that counts only "hits" will not reflect the true nature of the site's use. In addition, neither "hits" nor accesses translate directly into distinct users. Many Internet service providers, such as America OnLine, use "proxy" servers. Because of this, the Access Log will not accurately trace the number of users but, rather, reflect the number of accesses/"hits" made by the referring server.
 - A related issue is understanding the context of the server and presenting the data within that context. For example, this study concentrated on the use of the GILS Web server in the context of the EPA Web server. Readers cannot not, however, interpret the Web-based GILS record use as scant, moderate, or high without knowing the usage of each EPA server component. This study did not seek to provide that context.
- *Knowing what data to count.* Each Web server has different file naming conventions and methods of organization. For example, the EPA server used such file extensions as .txt to designate Web pages (as opposed to the more commonly used .htm or .html extensions). In order to accurately reflect the page accesses, the study team re-wrote the PERL scripts to count .txt extension files as accesses rather than hits. It is not clear to the study team, however, if these naming conventions hold throughout the entire server. Therefore, some accesses may actually be represented as hits in this study.
 - A related, and important issue, is that of internal versus external EPA server use. EPA has several IP domains that access the EPA server on a daily basis. The issue is the extent to which some of those accesses and/or hits are due to public requests for information. There is no current way, as of today, to gather such data. In the future, however, it may be worth identifying, isolating, and analyzing server use by a selection of EPA domain addresses that serve as public information offices, for example, to gain a greater sense of the EPA server's public service provision activities.
- *Selecting and/or developing appropriate analysis software.* Web server administrators need to plan for the analysis of Web server log files. The types of information about Web server use desired by those running the server should drive the selection and/or development process of log analysis software. Web administrators should not retrofit their log file analysis to the capabilities of the software.
- *Obtaining the cooperation of server administrators and Internet service providers.* Not all networked information providers run their own servers or have direct control over the Web server on which the Web-based services reside. As such, it is important to gain the cooperation of those individuals and/or entities that have direct control over the log files. The lack of such cooperation will have a negative impact on the ability to attain Web server usage data.

- *Preserving the privacy and confidentiality of server users.* In some cases, it is possible to trace directly back to a user, depending on the method of access a user has to a Web site. Web service providers need to develop policies as to how such data, if at all, will be used. This issue is particularly troublesome for public sector organizations, as such capabilities may violate privacy laws.
- *Educating server administrators and decision makers as to the benefits of log file analysis.* Log file analysis is just beginning to gain popularity. Server administrators and decision makers need to understand the types of data that log files can generate, the application of that data in an organizational setting, and the incorporation of such data into management activities.
- *Managing the log analysis process.* Gaining access to and analyzing Web server log files requires planning and coordination. To engage in log file analysis activities, there needs to be a delegation of responsibility for making the files available (on-site or remotely), performing the analysis (on-site or remotely), interpreting the analyzed data, and reporting the findings. Moreover, such analysis needs to be performed and reported on an ongoing and regular basis.
- *Presenting Web log statistics effectively on the Web itself.* Two issues require resolution: (1) the presentation of Web usage statistics on the Web; and (2) the means through which to display such statistics. Increasingly, users want to review Web server statistics of the sites they visit. This requires the presentation of those statistics by the site providers. Since Web usage statistics are in their infancy, however, little is known about appropriate ways in which to display such usage data and the purpose that is served in doing so.

These issues serve as a beginning point for Web server log analysis collection, presentation, and interpretation. Other issues exist, and still more will develop as Web services increase and log analysis techniques become more sophisticated.

7.0. MOVING FORWARD

Research into the analysis of Web server log files is limited. Web server administrators and decision makers are just beginning to understand the potential for systematic server usage data, and researchers are only just beginning to develop sophisticated analysis techniques. Key areas that require further exploration include the:

- *Ability to export files and analyze them in other formal statistics programs.* Current analysis techniques require specialized software and/or the development of specific analysis programs. There is a need to develop means through which log files can be imported and analyzed using off-the-shelf statistical analysis programs.
- *Understanding of log file data as user-based measures of Web services.* By performing Web log file analysis, server administrators and decision makers can begin to understand the path users take through a server, the problems users encounter during a session, and technology users use while navigating a site. Together, these are powerful data that can assist in the planning and design of Web-based services.
- *Cross-referencing log files.* This is an area of analysis that intends to cross-tabulate the various log files. For example, by cross-referencing the Error and Access Log files, one could know how many users, after receiving an error, stop surfing the site on which the error was received. To find this percentage a server administrator would use the domain name and time of the user who received an error (from the Error Log) and then look in the Access Log to see if that domain name shows up after the time of the error.
- *Creation of script files that can assess multiple types of log files.* While there are certain log file standards, not all Web server log files are exactly alike. Until such time as all log files are the same, the development of log analysis scripts will need to be able to accommodate multiple log file types so as to generate the same types of information regardless of file type.
- *Customization of script files.* Even if all Web servers generate log files that conform to certain standards, there will likely be differences in Web page and file naming conventions across servers. As such, script files will require modifications to meet the needs of specific log data.

- *Separation of internal versus external server traffic.* In order to determine the user community of the server, Web server administrators need to know who is accessing the server. Cursory analysis of the EPA log files indicates that a substantial proportion of the EPA Web server is generated by EPA IP addresses. It would be useful to know if EPA-generated Web server traffic and use differs than that of non-EPA users. Such data would assist the server administrators customize various portions of the server to more specifically meet the needs of various user groups.
- *Incorporation of log file analysis with other on-going electronic network assessment techniques.* The assessment and evaluation of electronic networks and network-based resources is increasing in scope and application (see Moen & McClure, 1997; Bertot & McClure, 1996; McClure & Lopata, 1995). Web log file analysis is a network-based assessment technique that is particularly useful when performed in conjunction with other on-going evaluation activities.

There is a need to resolve, minimally, these issues and move the ability to perform log file analysis forward. Log file data can provide user-based measures of Web-based resources if performed on a regular basis, incorporated into other electronic network assessment activities, and interpreted correctly.

This study presents a beginning point for Web log file assessment techniques. Researchers, server administrators, and decision makers are just now starting to understand the potential for Web log file analysis as part of a larger user-based measure of electronic resources. As the Federal government increases its provision of Web-based services for its citizens, agencies will need to develop, implement, and maintain an on-going assessment of Web-based activities through the analysis of Web server log files.

REFERENCES

Bertot, John Carlo & McClure, Charles R. (1996). *Sailor Assessment Final Report: Findings and Future Sailor Development*. Baltimore, MD: Division of Library Development and Services.

Christian, Eliot J. (1996, December). GILS: What is it? Where is it going? *D-Lib Magazine*. <<http://www.ukoln.ac.uk/dlib/dlib/december96/12christian.html>>.

Christian, Eliot J. (1994, May). *The Government information locator service (GILS): Report to the information infrastructure task force*. <<http://www.usgs.gov/gils/gilsdoc.html>>.

McClure, Charles R., and Lopata, Cynthia. (1996). *Assessing the Academic Networked Environment: Strategies and Options*. Washington, D.C.: Coalition for Networked Information.

McClure, Charles R., Moen, William and Ryan, Joe. (1992). Design for an Internet-Based Government-Wide Information Locator System. *Electronic Networking: Research, Applications, and Policy*, 2, 6-37.

Moen, William E. & McClure, Charles R. (1997, January 14). Multi-Method Approach for Evaluating Complex Networked Information Services: Report from an Evaluation Study of the Government Information Locator Service (GILS). Technical Paper Submission: ACM Digital Libraries '97. [Paper Submitted].

Moen, William E. & McClure, Charles R. (1996, August). *Technical Proposal: An Evaluation of the Federal Government's Implementation of the Government Information Locator Service (GILS)*. Denton, TX: University of North Texas, School of Library and Information Sciences. <<http://www-lan.unt.edu/slis/research/gilseval/gilseval.htm>>.

Moen, William E. & McClure, Charles R. (1994). *The Government Information Locator Service (GILS): Expanding Research and Development on the ANSI/NISO Z39.50 Information Retrieval Standard, Final Report*. Prepared for the United States Geological Survey and the Interagency Working Group on Data Management for Global Change, Washington, DC [USGS, Cooperative Agreement No., 1434-93-A-1182]. Bethesda, MD: NISO.

National Institute of Standards and Technology, Computer Systems Laboratory. (1994, December 7). *Federal Information Processing Standards Publications 192, Application Profile for the Government Information Locator Service (GILS)*. Gaithersburg, MD: U.S. Department of Commerce, National Institute of Standards and Technology. <<http://www.dtic.dla.mil/gils/documents/naradoc/fip192.html>>.

Network Wizards. (1997). *Internet Domain Survey, January 1997*. <<http://www.nw.com>>.

Noonan, Dana. (1996). Making Sense of Web Usage Statistics. *The PIPER Letter: Users and Usage*. <<http://www.piperinfo.com/piper/pl01/usage.html>>.

Novak and Hoffman. (1996). New Metrics for New Media: Toward the Development of Web Measurement Standards. <<http://www2000.ogsm.vanderbilt.edu/novak/web.standards/webstand.html>>.

Rubin, Jeff. (1996). Log Analysis - A Brief Overview. <<http://headcase.syr.edu/text/logs.html>>.

Stout, Rick. (1997). *Web Site Stats: Tracking Hits and Analyzing Traffic*. Osborne McGraw Hill: Berkley, CA.

APPENDIX F
The Role of GILS Metadata in
Networked Information Discovery and Retrieval

APPENDIX F

The Role of GILS Metadata in Networked Information Discovery and Retrieval

GILS records are an early innovation in the use of metadata in the networked environment to describe and point to selected information objects (whether or not those objects were available electronically). GILS has not been the only Federal initiative concerned with metadata. For example, the Federal Geographic Standards Committee (FGDC) worked for several years to develop a metadata standard that would include appropriate elements for describing geospatially-referenced information (Mangan, 1995). The efforts of GILS and FGDC, however, were simply the precursor to activities by others in the networked environment to devise a way of describing network-accessible objects (whether documents, images, multi-media objects, etc.) so that they could be discovered, identified, and accessed (see for example the work on the Dublin Core metadata elements in Weibel, et al., 1995).

One of the meanings of the term metadata in GILS refers to the actual set of data elements that comprise a GILS record. Thus, the terms "metadata record" and "GILS record" are interchangeable, or, more explicitly, GILS records are metadata records. The data elements defined for use in GILS records constitute a metadata scheme. The scheme includes specific names of elements, definitions for the elements, and their structure. The GILS elements provide a standardized way of representing information objects—whether they be online or not, a low or high level of aggregation, etc.

The term metadata can also refer to the "locators" that GILS records may describe; GILS records may describe information resources that contain metadata records. For example, if there is a GILS record for an index or catalog of agency publications, that catalog may be considered "metadata" in the sense that entries in the catalog provide data about and serve as pointers to information resources.

Descriptive metadata contained in GILS records are a foundation for several processes that include discovering, locating, and accessing information. Users can make initial relevance judgments about a resource simply by examining the metadata record that describes the resource. Standardized data elements and content in the record can improve search and retrieval. GILS records can be also support machine-processing such as done by Web robots.

APPENDIX G
Characteristics of Successful Objectives

APPENDIX G

Characteristics of Successful Objectives

Identifying and articulating appropriate objectives for the "refocused GILS" is essential. One approach, adapted from Total Quality Management, identifies five aspects in articulating objectives: Specific, Measurable, Accountable, Realistic, Time-Phased. The deployment of TQM and its associated requirements for expression of organizational mission statements and policy has given rise to this mnemonic device called "SMART."

The aspects of SMART objectives outlined below can be contrasted with current language of goals, objectives, purposes, and expectations in GILS policy. The GILS II initiative can refocus the goals and scope for GILS, but that is just the first step. Policy, as articulated by OMB, then needs to be translated in SMART objectives.

- **Specific:** This characteristic provides focus, or an ability to visualize a clear outcome or planned state. Definitions must be operationalized and understood (but not necessarily agreed upon) by everyone that is expected to participate. Examples of nonspecificity in OMB Bulletin 95-01 that have hindered GILS implementations to-date include:
 - "public information resources"
 - "automated information system"
 - "...all departments and agencies in the Executive Branch" [and] "Independent regulatory commissions and agencies"

The lack of specificity in just these three items alone limit one's ability to envision GILS outcome or a state of requirements satisfaction. The record aggregation issue also is consequence of nonspecificity, as is the confusion surrounding the concept of "US Federal core" locator records. OMB Bulletin 95-01, FIPS 192, and the NARA *Guidelines* fail to specify precisely what these are and what purpose they serve beyond that of "non-Core" items.

- **Measurable:** If progress toward an objective can't be measured, it won't be accomplished. It goes hand-in-hand with specificity; if something can't be visualized, it can't be broken down into recognizable/countable units. The objectives of GILS have not been measurable. For example:
 - "Assistance" in obtaining the information and "help" the public and agencies are goals not easily measurable (or at least the instruments are not available to, trusted by, or usable by the implementors). Such goals need to be described in terms of measurable criteria.
 - "improve agency electronic records management practices" [and] "agencies' abilities to carry out their records management responsibilities and to respond to Freedom of Information Act requests"—what are the benchmarks? "Improvement" implies they are known.
 - "...agencies should inventory their existing holdings and institute adequate (how to measure "adequate"?) information management practices. To the extent practicable (how is practicality measured?), agency GILS should contain automated links to underlying databases to permit direct access to information..."
 - "Reduce (to what level?) the information collection burden on the public by making existing (at what point in time?) information more (how much more?) readily available for sharing among agencies.
- **Accountable** – Policy states that the "head" of each agency is accountable. But for what? The "head" of any organization is always accountable for everything. With the lack of agency management support identified through the study, agencies may have missed the boat here by not specifying responsibilities and authorities in the chain of command. Also, "Independent regulatory commissions and agencies are requested to comply" automatically abrogates accountability.

- **Realistic** -- One tenet of realistic modeling is precedent, which agencies may be lacking in areas of IRM and RM in the evolving networked environment. "Realistic" can be contrasted with "reasonable." The original vision of GILS might have been reasonable, but was it *realistic*?
 - Is centralization or decentralization or some hybrid model realistic?
 - Information dissemination product means "any book, paper, map, machine-readable material, audiovisual production, or other documentary material, regardless of physical form or characteristic, disseminated by an agency to the public." This scope ("any") may be reasonable, but *realistic*?
- **Time-Phased** -- The key here is the concept of "sufficiently frequent cause for celebration or reflection." OMB Bulletin 95-01 has several "phases" for the objectives of GILS, but without co-incident S-M-A and R. During the study, no participants suggested that any celebration or reflection occurred on making an OMB Bulletin 95-01 deadline. Objectives need to be associated with specific time frames for their accomplishment.

An example of a "SMART" government-wide objective for GILS implementation might have been:

By January 1, 1996, the manager charged with agency GILS implementation in each Cabinet Department and Executive Agency listed in the *1996-97 Government Manual* shall mount on the agency web site, or on GPO Access, a metadata record comprising Title, Abstract, Order Process, and Point of Contact for their 10 most frequently-requested printed publications.

Implicit in this is the idea that, in an empowered culture, the S-M-A-R-Tness gets stronger with every link in the organizational chain—i.e., "granules" of responsibility become apparent. In the above example, the Public Information Officer, for example, would recognize that his/her contribution is to identify the popular resources (and that if he/she doesn't already collect the data that he/she had better have it by a date negotiated by the team); the webmaster allocates *n* bytes; etc.

The lack of S-M-A-R-T objectives for GILS could be simply a reflection that at the outset the philosophical mandate for GILS was unclear. *It will be necessary to pinpoint as closely as possible who wants what and why from the refocused GILS.* That understanding, and necessary buy-in by the agencies, could ease the development of S-M-A-R-T objectives for the "how." Further, the accomplishment of the objectives should lead to tangible benefits to the agencies.

APPENDIX H
GILS Record Content and Display Variation

APPENDIX H

GILS Record Content and Display Variation

The two GILS records in this appendix reflect dramatic variation in content and display. In the first instance, most of the GILS data elements appear but with the value NONE. The second instance is rich in descriptive detail. These records are reproduced here exactly as a user would discover them (e.g., file format, layout, etc.).

Title: INDIVIDUALS WITH DEFERRED VESTED PENSION BENEFITS

Acronym:

Originator:

Department/Agency Name: Social Security Administration

Major Organizational Subdivision:

Minor Organization Subdivision:

Name of Unit:

Controlled Vocabulary: NONE

Local Subject Index:

Local Subject Term: US Federal GILS

Abstract: NONE

Purpose: This file records the nature and form of the pension benefit, the name of the plan, employer identification number (EIN) of the plan's sponsor, plan number, name and address of the plan administrator, and the EIN of the plan administrator.

Agency Program: NONE

Availability: NONE

Sources of Data: NONE

Methodology: NONE

Access Constraints: NONE

Use Constraints: NONE

Point of Contact for Further Information:

Name: Josephine T. Iampieri

Organization: Office of Central Records Operations

Street Address: 12th Floor, MetWest Tower, 300 N. Greene St.

City: Baltimore

State: MD

Zip Code: 21201

Country: US

Network Address:

Hours of Service:

Telephone: 410-966-8711

Fax:

Supplemental Information:

File Code: ERN-13.00.00

SAC Code: S2B

Cross Reference: NONE

Schedule Number: NONE

Control Identifier: NONE

Record Source: NONE

Original Control Identifier: NONE

Disposition:

FRC Storage Authorized: N

Disposition Authority:

Description:

Date of Last Modification: 19951212

Title: Department of Labor

Acronym:

Superintendent of Documents Class Number Stem: L

Local Subject Index:

Local Subject Term: U.S. Federal GILS

Local Subject Term: Accident Prevention

Local Subject Term: Accidents

Local Subject Term: Business

Local Subject Term: Career Education

Local Subject Term: Civil Rights and Equal Opportunity

Local Subject Term: Collective Bargaining

Local Subject Term: Contracts

Local Subject Term: Cost of Living

Local Subject Term: Counseling

Local Subject Term: Discrimination

Local Subject Term: Economic Policy

Local Subject Term: Education

Local Subject Term: Employment and Occupations

Local Subject Term: Equal Opportunity

Local Subject Term: Factories

Local Subject Term: Handicapped

Local Subject Term: Health

Local Subject Term: Health Care

Local Subject Term: Industrial Safety

Local Subject Term: Industry

Local Subject Term: Insurance

Local Subject Term: International Trade

Local Subject Term: Labor-Management Relations

Local Subject Term: Management

Local Subject Term: Minorities

Local Subject Term: Occupational Outlook Handbook

Local Subject Term: Occupational Safety and Health

Local Subject Term: Occupations

Local Subject Term: Personnel Management

Local Subject Term: Physical Fitness

Local Subject Term: Physically Challenged

Local Subject Term: Recreation

Local Subject Term: Retirement

Local Subject Term: Safety

Local Subject Term: Social Security

Local Subject Term: Unions

Local Subject Term: Veterans

Local Subject Term: Vital and Health Statistics

Local Subject Term: Vocational Education

Local Subject Term: Vocational Guidance

Local Subject Term: Wages

Local Subject Term: Women

Local Subject Term: Workers' Compensation

Abstract: "The purpose of the Department of Labor is to foster, promote, and develop the welfare of the wage earners of the United States, to improve their working conditions, and to advance their opportunities for profitable employment. In carrying out this mission, the Department administers a variety of Federal labor laws guaranteeing workers' rights to safe and healthful working conditions, a minimum hourly wage and overtime pay, freedom from employment discrimination, unemployment insurance, and workers' compensation. The Department also protects workers' pension rights; provides for job training programs; helps workers find jobs; works to strengthen free collective bargaining; and keeps track of changes in employment, prices, and other national economic measurements. As the Department seeks to assist all Americans who need and want to work, special efforts are made to meet the unique job market problems of older workers, youths, minority group members, women, the handicapped, and other groups."

Purpose: Selected publications of this and other United States government agencies are available through the Federal Depository Library Program (FDLP). The index to the publications available through the FDLP is the Monthly Catalog of United States Government Publications (MOCAT), which also includes ordering information for those publications available for sale. Many publications are available to the public in the approximately 1400 Federal depository libraries throughout the United States. See also data in this record for Access Constraints and User Constraints, where such limitations on this agency's publications are appropriate.

Agency Program: "The Department of Labor (DOL), the ninth executive department, was created by act of March 4, 1913 (29 U.S.C. 551). A Bureau of Labor was first created by Congress in 1884 under the Interior Department. The Bureau of Labor later became independent as a Department of Labor without executive rank. It again returned to bureau status in the Department of Commerce and Labor, which was created by act of February 14, 1903 (15 U.S.C. 1501)."

Availability: "The Office of Public Affairs distributes a brochure entitled 'Department of Labor,' which describes the activities of the major agencies within the Department, and 'Publications of the Department of Labor,' a subject listing of publications available from the Department."

Distributor:

Distributor Name: Office of Public Affairs, Department of Labor

Distributor Street Address: Room S-1032, 200 Constitution Avenue, N.W.

Distributor City: Washington, D.C.

Distributor Zip Code: 20210

Distributor Network Address:

Distributor Telephone: 202-219-7316

Access Constraints: No access constraints apply to information disseminated through the Federal Depository Library Program. In other cases, some access constraints may be applied by the agency to assure the protection of privacy or intellectual property, and any other special restrictions or limitations on obtaining the information resource.

Use Constraints: No constraints apply to the use of information disseminated through the Federal Depository Library Program. In other cases, some use constraints may be applied by the agency to assure the protection of privacy or intellectual property, and any other special restrictions or limitations on using the information resource.

Cross Reference:

Cross Reference Title: Monthly Catalog of U.S. Government Publications

Cross Reference Linkage: http://www.access.gpo.gov/su_docs/dpos/adpos400.html

Cross Reference Type: HTTP

Cross Reference:

Cross Reference Title: U.S. Department of Labor [Home Page]

Cross Reference Linkage: <http://www.dol.gov>

Cross Reference Type: HTTP

Point of Contact for Further Information:

Organization: Office of Public Affairs, Department of Labor

Street Address: Room S-1032, 200 Constitution Avenue, N.W.

City: Washington

State: D.C.

Zip Code: 20210

Telephone: 202-219-7316

Supplemental Information: Source of information for Local Subject Index is Subject Bibliography Index. Most of the other information in this record is extracted from the U.S. Government Manual.

Control Identifier: GPO/SOD/LPS-0027

Record Source:

Department/Agency Name: United States Government Printing Office

Major Organizational Subdivision: Superintendent of Documents

Minor Organizational Subdivision: Library Programs Service

Name of Unit: Cataloging Branch

Internet Contact: dldgpo@access.digex.net

Telephone: 202-512-1141

Date of Last Modification: 19960515

APPENDIX I
Preliminary Report of the Canadian GILS Subgroup
And GILS Pilot Project
(Reprinted with Permission)

Government Information Locator Service

(GILS)

Preliminary Report
of the
Canadian GILS Subgroup
and
GILS Pilot Project

February 24, 1997

TABLE OF CONTENTS

EXECUTIVE SUMMARY	i
BACKGROUND	1
1. WHAT IS GILS?	1
2. WHY INFORMATION RESOURCE DESCRIPTION?	2
2.1 BENEFITS OF INFORMATION RESOURCE DESCRIPTION	2
2.2 WHY STANDARD METADATA?	3
2.3 INTERNET ACTION GROUP ON DOCUMENT IDENTIFICATION	3
3. CANADIAN GILS ACTIVITIES	4
3.1 GILS SUBGROUP	4
3.2. EXTENDING THE GILS STANDARD - GSG CONTRIBUTION	4
3.3 DEVELOPMENT OF THE CANADIAN GILS GUIDELINES	5
3.4 SPECIFICATION OF A GILS RECORD SYNTAX	5
3.6 PRESENTATIONS AND PARTICIPANT TRAINING	6
4. GILS IMPLEMENTATION	6
4.1 GOVERNMENT INFORMATION FINDER TECHNOLOGY: TARGET ARCHITECTURE	6
4.2 RECORD CREATION OPTIONS	7
4.3 EVOLUTION OF DEPARTMENTAL GILS SITES	10
4.4 U.S. IMPLEMENTATION EXPERIENCE	11
4.5 OTHER GOVERNMENTAL GILS INITIATIVES	13
5. GILS PILOT PROJECT - PHASE I	14
5.1 PROJECT OBJECTIVES	14
5.2 PROJECT PARTICIPANTS	14
5.3 EVALUATION EXERCISE	15
5.4 SUMMARY AND INTERPRETATION OF RESPONSES TO RECORD CREATION FACILITIES	15
5.5 SUMMARY AND INTERPRETATION OF RESPONSES TO SEARCH FACILITIES	17
6. PHASE I FOLLOW-UP AND DOWNSTREAM DEVELOPMENT	18
6.1 LINKAGE WITH OTHER INFORMATION DELIVERY SERVICES	18
6.2 THESAURUS SUPPORT	18
6.3 REGISTRATION AND ADMINISTRATION OF INFORMATION RESOURCE IDENTIFIERS	19
6.4 GUIDELINES FOR DEPARTMENTAL SERVER IMPLEMENTATION	19
6.5 SECURITY AND AUTHENTICITY OF GILS RECORDS	19
6.6 ENHANCEMENTS TO THE CENTRAL SEARCH FACILITY	20
6.7 MAINTAINING THE CANADIAN GILS GUIDELINES	20
6.8 DEVELOPMENT OF RESOURCE DESCRIPTION GUIDELINES	20
6.9 ADDITIONAL SUPPORT FOR ARCHIVING FUNCTION	20
6.10 MAINTENANCE OF CENTRAL GILS DATABASE	21
6.11 LINK TO DIRECTORY APPLICATIONS	21
6.12 COMMUNICATIONS PLAN AND TRAINING SUPPORT	21
6.13 INTEGRATING CATALOGUING AND GILS RECORDS	22
6.14 INTEGRATING GILS WITH INFORMATION WORKFLOW MANAGEMENT	22
7. RECOMMENDATIONS	22

7.1 POLICY FACTORS.....	23
7.2 OPERATIONAL FACTORS	23
7.3 TECHNICAL FACTORS	24
APPENDIX A - MEMBERS OF THE GILS SUBGROUP.....	26
APPENDIX B - QUESTIONNAIRE FOR RECORD CREATORS.....	27
B.1 PROFILE OF GILS RECORD CREATOR	27
B.2 TIME SPENT CREATING RECORDS	27
B.3 AVAILABILITY OF INFORMATION REQUIRED.....	28
B.4 WORK ROUTINES	32
B.5 RECORD CREATION TOOL.....	32
B.6 THE CANADIAN GILS GUIDELINES	33
APPENDIX C - QUESTIONNAIRE FOR SEARCH PARTICIPANTS	34
C.1 GILS SEARCH FACILITIES	34
C.2 GENERAL QUESTIONS	36
APPENDIX D - BIBLIOGRAPHY	37
APPENDIX E - MAPPING GILS AND INFOSOURCE DATA ELEMENTS	39

Executive Summary

This preliminary report of the Canadian GILS Subgroup and the GILS Pilot Project summarizes the activities and progress that has been made to implement a Canadian Government Information Locator Service (GILS).

In August 1995, Treasury Board recognized a need to establish a primary Government of Canada Internet site and requested the Government Telecommunications and Informatics Services (GTIS) to develop and maintain gateway services to government information.

Based on a U.S. standard, adapted to meet Canadian government needs, GILS provides users with the means of finding government information located in local and remote systems. It offers a standard way of describing government information holdings. These standardized descriptions may in turn be used by automated systems to improve the precision of the information retrieval process and also to assist government departments in managing their information holdings. GILS is based on standards for information retrieval and interchange. Thus it can be implemented on any systems hardware/software platform that is connected to the Internet or an Intranet.

Section 3 provides a very brief summary of Canadian GILS activities. It describes extensions to the U.S. standard; highlights development of the Canadian GILS Guidelines; identifies valid GILS record formats, and indicates initial promotional and training efforts provided by the subgroup members.

Section 4 describes GILS implementation experience in Canada and abroad. It summarizes development work done by GTIS; gives departmental options for GILS record creation, and; notes the opportunity for distributed GILS database deployment. The U.S. experience with GILS implementation is presented and other GILS initiatives noted to position the Canadian government efforts in a global context.

Section 5 gives the Phase 1 project objectives; identifies the participating departments; summarizes user responses to two pilot questionnaires, and; interprets the initial user feedback on the GILS record and information retrieval facilities.

Section 6 identifies a significant number of issues and requirements that will need to be addressed in a Phase 1 follow-up and through downstream development on a government-wide and departmental basis.

The report concludes with a set of recommendations concerning policy, operational and technical factors related to government-wide GILS deployment. It recommends government-wide commitment to GILS and identifies the need for co-ordination within and across departments, staff training and development and technical leadership.

Appendices A through E provide further details on items mentioned in the report.

Background

In March 1995, the Treasury Board Information Management Sub-Committee (TIMS) approved the government-wide Internet strategy (For details see: http://canada.gc.ca/programs/guide/1_1_4e.html). The strategy recognized the need for a corporate-wide federal presence on the Internet and a focal point for single-window access through the Internet to services and information available from federal institutions.

In August 1995, Treasury Board approved the selection and mandate of Government Telecommunications and Informatics Services (GTIS) to host the primary Government of Canada Internet site (Canada site). For details concerning the mandate see: <http://www.pwgsc.gc.ca:80/homepage/text/g1gtis-e.html>. Treasury Board also recognized the need for improved location, search and retrieval capabilities. Specifically GTIS was asked to:

- incorporate the technology and information of the Open Government Pilot and the Champlain search and retrieval software developed by Industry Canada with the Government Information Finder Technology (GIFT) developed by GTIS, and;
- develop and maintain Internet gateway services including government information location and retrieval capabilities.

GILS is the standard and service strategy, that was adopted by GTIS, to implement GIFT and to provide the gateway location and retrieval capabilities. This strategy was based on recommendations and implementation advice provided by the GILS Subgroup, an inter-department working group established under the aegis of the Government Standards Program.

GILS is also the standard recommended in the Government of Canada Internet Guide for internal departmental use and for sharing metadata across departments.

1. What is GILS?

The Government Information Locator Service (GILS) is a computer platform independent system for locating government information in a decentralized collection of databases. GILS systems or locators are made up of searchable databases of GILS records which indicate what information is available, where it is located and how it may be accessed or acquired. A GILS record is not the information itself, but a standards-compliant description and a pointer to an information resource. GILS records can describe a collection, a service, a system, a Web site, a publication or an individual electronic document. They can contain a direct link (Uniform Resource Locator or URL) to a networked information resource. They can also describe how to obtain information that is not available on an electronic network such as the Internet or a departmental Intranet.

GILS originated in the United States. U.S. federal government agencies were required by law to implement this government-wide service beginning January 1996.

This service provides users with a means of finding government information, located in local and remote systems. Users find information by formulating system independent queries at the desktop and transmitting these to a remote database containing GILS records. The queries are presented to the remote database in an international standard language for information retrieval called Z39.50. It provides rules and procedures for the exchange of information between two systems independent of what hardware or software those systems run on. Z39.50 allows users to search one or more databases and to receive a consolidated set of responses to each search query. Recognizing that Z39.50 implementations will not be prevalent at the user's desktop, government information providers typically provide a Web/Z39.50 gateway to this information service. These gateways include support for the Internet-HTTP standard, supported by World Wide Web browsers, and thereby provide access to government information for anyone who has a Web browser.

2. Why Information Resource Description?

2.1 Benefits of Information Resource Description

The power of full text searching, as demonstrated in the Champlain project and other Internet-based implementations, provides users with an ability to identify vast amounts of information located in various sources. The identified resources typically include many irrelevant and duplicate items. Thus the user is left with the choice of reformulating the search in anticipation of improved relevancy or sifting through voluminous amounts of irrelevant and duplicated references.

Web crawlers and database search engines can operate much more efficiently and precisely if the information that they index and search is described in a precise and compact manner. Such precise and compact description of the content, structure and associated features of an information resource is referred to as metadata. Analogous terms such as uniform resource identifiers (URIs), uniform resource locators (URLs), and uniform resource names (URNs) are used within the Internet community in recognition of the need to enhance information access by standardizing information resource descriptions. Rather than searching an index of the entire text or some arbitrary portion of each document or information resource description the precision and relevance of search results could be improved by restricting the search to the metadata. Metadata identifies specific elements of an information resource such as the title, the author, the subject, the creation date, etc. In so doing it can make explicit information that cannot be readily deduced from the information resource itself (e.g. originator, language of resource, physical characteristics of the information container or medium, etc.). Metadata is essential in order to document information resources, to indicate their structure, the format of elements within the resources; what software must be used to access them, etc. Proper description of information resources enables search engines to focus, and

optionally weight, the search terms. It also enables a user to determine more readily and accurately the usefulness of an information resource prior to downloading it to the user's site.

Metadata can also support information resource management. For example, review dates within a resource description can trigger human review or automatic update of outdated information resources. Metadata can include information about security, authentication, or preservation of an information resource and can support version control. Metadata can also support service objectives, for example, identification of new resources based on date to automatically create a What's New page on a Web site. Metadata can be used to better link information to a user's needs, for example, deliver information only in the industrial sector or geographic area that match the user's interests.

2.2 Why Standard Metadata?

The primary objective in standardizing the metadata for government information resources is to facilitate user access to this information as envisaged by various single window scenarios. Basically users should be presented with a single view of the available government information resources. This view would hide the technicalities of information retrieval, indexing, display and related characteristics of individual government systems. The user would be left with the impression of accessing a single government-wide system.

The first step towards realizing this vision was taken in the GILS pilot project by reaching agreement on a uniform means of describing government information resources, the GILS record content and syntax. The content specifies the descriptive data (i.e. metadata) and the syntax prescribes the format that supports intersystem record exchange and processing. Actually there are separate syntaxes for information retrieval and for record exchange as noted in section 3.4.

GILS attribute sets and the GILS schema are registered objects (i.e. are globally identified and designated for use with Z39.50). The GILS Core Element Set with the additions recommended by the Canadian GILS Subgroup and others has proven to be extensible and flexible enough to describe a wide variety of information resources to various levels of detail.

2.3 Internet Action Group on Document Identification

The first federal government Internet conference, held in March 1994, concluded that "document identification" was an essential component for supporting access to government information. An Internet Action Group on Document Identification, consisting of federal librarians and a representative from the Depository Service Program, was established to explore the challenges associated with finding federal government information on the Internet. This group concluded that finding government documents on the *Internet* was indeed difficult and that it was hard to verify that a located Internet

document did indeed come from the federal government. The apparent solution was provision of "metadata" for government publications using one of several evolving metadata standards examined by the Action Group.

3. Canadian GILS Activities

3.1 GILS Subgroup

To formally address the perceived need for a government-wide metadata standard, the Treasury Board Internet Advisory Committee and the Electronic Document Standards Working Group (EDSWG) agreed that a GILS Subgroup (GSG) should be established in November 1995 within the Government Standards Program.

The mandate of the GILS Subgroup is to prepare a draft Government standard for describing federal government information resources in order to:

- organize and manage information resources in a consistent and systematic manner;
- facilitate the implementation of precision searching on the Internet or other wide area networks, and;
- provide improved service to end users by providing multiple access points to information resources through use of metadata descriptions.

3.2 Extending the GILS Standard - GSG Contribution

Early in its deliberations, the GILS Subgroup determined that the U.S. GILS standard required adaptation before it could meet Canadian government needs. For example, extensions and modifications were needed to:

- indicate and describe information resources in both official languages;
- describe individual documents and publications;
- identify classified and restricted information resources, and;
- develop a syntax to support record interchange.

An effective liaison was maintained throughout the spring of 1996 that allowed the GILS Subgroup to propose revisions and additions to the U.S. GILS standard (also called the GILS Profile). Following due assessment and consideration the U.S. GILS committee incorporated the various Subgroup requirements into version 2 of the GILS Profile finalized in October 1996.

This collaboration helped to "internationalize" the original version of the U.S. standard. For example, the indication of mandatory and optional elements was removed to allow national implementations to determine which elements should be mandatory.

This revision enables the element for "language of resource" to be mandatory in the Government of Canada but remain optional for U.S. government agencies.

3.3 Development of the Canadian GILS Guidelines

The Canadian GILS Guidelines, prepared by Fay Turner of the National Library, provides definitions and examples of individual elements and identifies sources of information in certain cases. Based on the GILS Guidelines, developed by the U.S. National Archives and Records Administration, the Canadian version supports federal government practices such as the use of Federal Identity Program symbols to identify the originating department. Other examples supplied by the Subgroup members included clarifications such items as the "schedule number" which is to be provided in accordance with National Archives of Canada guidelines.

Recommendations of the GILS Subgroup have been incorporated into the new edition of the Government of Canada Internet Guide. The Internet Guide recognizes that: "Users of federal government information on the Internet need to know that they can locate the information they want, and that the information they locate is up to date, accurate and authentic. Chapter 2, *Laying the Groundwork*, includes a section on providing metadata on information products. Chapter 3, *Building the Site*, includes an introduction to GILS and precision searching.

3.4 Specification of a GILS Record Syntax

GILS records may be transferred between two systems for two distinct purposes: 1) bulk record transfer to populate databases, and 2) retrieval of individual records by search clients. The interchange format for these two purposes is different.

3.4.1 Record Transfer and the GILS Interchange Format

One of the objectives of the Canadian GILS Pilot Project is to demonstrate the exchange of large files of GILS records between GILS systems; in particular, between record creator's systems and the GTIS central repository of GILS records. Records transferred for the purpose of replicating data and populating GILS databases should be encoded as SGML documents. This is a requirement for the transfer of records to the GTIS repository. The encoding is specified in the GILS document type definition (DTD) which is available from the GILS Web site: <http://gils.gc.ca/>

3.4.2 Information Retrieval and GILS Interchange Formats

The interchange format for a GILS record transferred to a client system in response to a Z39.50 query of a remote database will depend on both the formats supported by the client and the formats supported by the GILS server. For a description of the available formats, also known as transfer syntaxes, see the Application Profile for the Government Information Locator Service (GILS), Version 2. The formats supported for the transfer of GILS records via the Z39.50 protocol are:

- USMARC format for machine-readable cataloguing information
- Generic Record Syntax (GRS)
- Simple Unstructured Text Record Syntax (SUTRS)
- Formats supported via the HTTP protocol: HTML and SGML

3.6 Presentations and Participant Training

Regular presentations on the work of the GILS Subgroup were made to the Electronic Document Standards Working Group and the Treasury Board Internet Advisory Committee. GILS was also covered in a presentation on metadata to the *Library Information Exchange Forum* (LIEF) in Spring 1996. The GILS Pilot Project was officially launched on October 30, 1996 with an Information Session for all interested departments. A half day workshop on GILS was held as part of the *Government on the Net 96* conference in November 1996. GILS Subgroup members have made various presentations to their own departments and to other departments.

To date no formal GILS training sessions such as those offered by NARA in the U.S. have been held. Feedback provided at various presentations and through the Pilot Project make it clear that a substantial amount of training is needed before the objectives, information and technical aspects of GILS can be fully appreciated and quality GILS records can be created.

4. GILS Implementation

Development of a government-wide service based on the GILS standard was driven by the resources and expertise made available through the Architect Program at the Government Telecommunications and Informatics Services (GTIS). These development efforts were based on the collective advice provided by the GILS Subgroup and supplemented with contracted expertise in specific areas such as standards for information retrieval (i.e. Z39.50), interchange formats (i.e. SUTRS) and document definition (i.e. SGML).

The following sections highlight the phase one implementations that have been achieved thus far.

4.1 Government Information Finder Technology: Target Architecture

To test the GILS standard and to demonstrate the service capabilities, the GTIS design team led by Oliver Javanpour, selected a system configuration based on the client server architecture associated with Z39.50 and other information retrieval standards. As illustrated in Figure 1, this implementation permits a user to post searches to a central database of GILS records and to retrieve records describing information resources (i.e. documents).

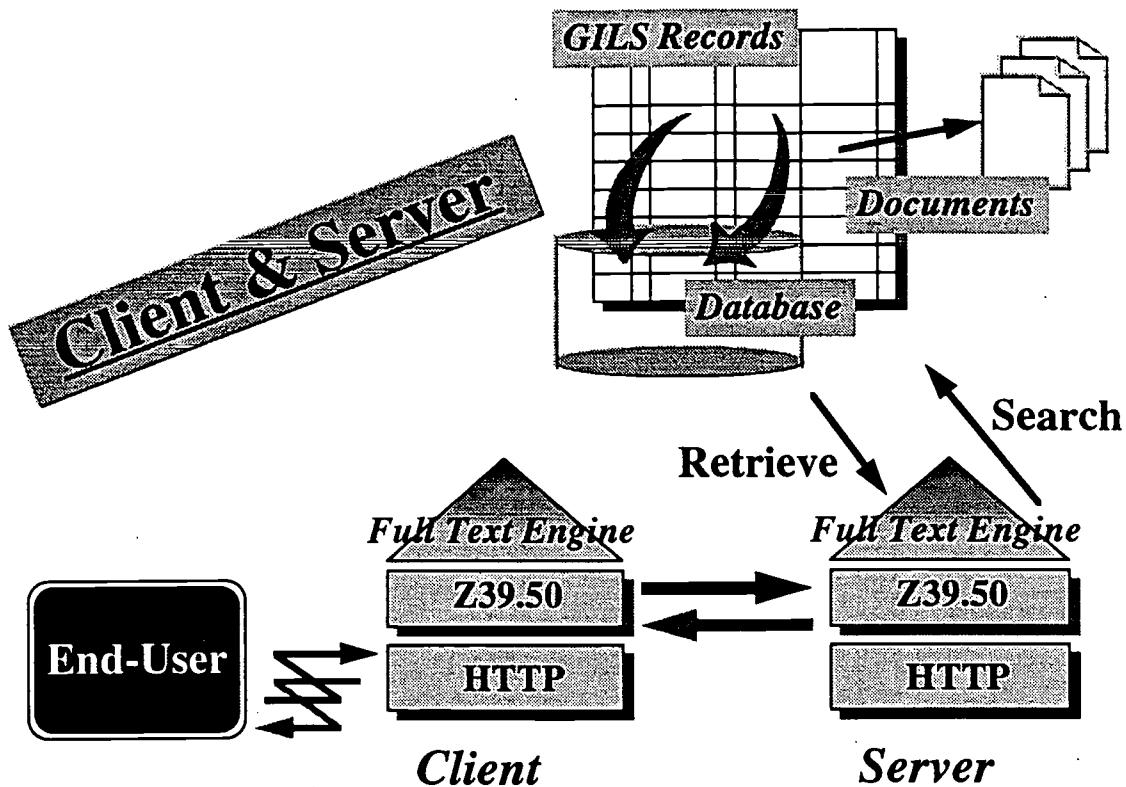


Figure 1 - GILS Pilot Search Facilities at GTIS Site

As illustrated above, a user has the option of using the http facilities available in Internet browsers or alternatively the client capabilities provided by commercial Z39.50 software.

4.2 Record Creation Options

Participating departments can create and interchange SGML-encoded records using one of three options. The best option for each participant will be determined, to a large extent, by departmental preferences and the existing information management infrastructure. The available record creation and contribution options include:

- Real-time record creation and contribution to the central repository using a WEB-based application tool developed for the project by GTIS. For details of the GILS Record Creation Tool see: <http://gils.gc.ca>;
- Scheduled batch conversion and contribution of existing information resource descriptions;
- Creation and contribution of GILS records from an SGML document environment. For details see Creating GILS Records in an SGML Environment available from the GILS Web site.

4.2.1 WEB Application Tool

To support record creation for the GILS pilot GTIS developed a Web-based tool for departments to capture the descriptive information and post it to the GILS database. This tool presents the record creator with a multi-page HTML form for recording the individual GILS elements and subelements. Mandatory information is highlighted and the record creator can supply as many of the optional elements as applicable, available or desirable. Certain elements are unique to specific information resources (e.g. an information resource isn't always linked to an agency program) or descriptive details may be missing at record creation time.

To reduce the record creation overhead, certain elements are filled according to default values established for each record creator and government department. These defaults are set when the creator registers with the GILS system; others are set as records are created and can be selected later from a pick list. Database maintenance functions have also been provided to enable record modification and deletion by record owners. Copies of newly created records are returned to the source department as e-mail attachments in the SGML interchange format and HTML.

4.2.2 Conversion of Existing Department Records to GILS

A data conversion tool was developed by GTIS to assist transformation of descriptive data in departmental records to the content and syntax specified in the Canadian GILS Guidelines. This tool was designed to support record interchange between existing departmental systems and the central GILS facility. Before this capability can be invoked each participating department must map the content of its descriptive information to the corresponding GILS record elements and values. Decisions and procedures for the periodic exchange and execution of the actual conversion are then established between GTIS and departmental support staff on a case by case basis.

4.2.2.1 Overview and Experience with the GILS Data Conversion Tool

GILS records are stored in an SGML format at the GTIS site. The data conversion tool was designed to provide government departments with a convenient method of converting existing descriptive records to a GILS SGML format. The conversion tool includes a data mapping (field name matching) facility to identify and match similar fields types which are then converted.

The process involves 3 steps. Initially the data must be assessed for matching field types and characteristics. Secondly the data must be queried and exported from its original database source to a text readable format. This option is available or can be developed for any database. Thirdly, the data must be processed through the

conversion utility. Following conversion of the data, it must be transferred to the GILS site through one of several options including e-mail and a web interface.

Several different record conversions were tested successfully and procedures are being explored for on-going operation with two source databases. Environment Canada provided a subset of its database which was queried and converted successfully. Conversion of a subset of the InfoSource source data encountered additional complexities associated with multiple record types.

Difficulties will arise whenever departments lack convenient access to internal expertise needed to extract the correct subset of the data from the database prior to conversion. If departments are to use the GILS conversion utility, some knowledge and technical expertise is needed in the existing departmental database system. This initial challenge can be solved through a one time effort by the departmental database administrator.

Due effort will have to be devoted to resolving complexities that can arise during the conversion process. For example, aligning of selection lists in the source application with values in corresponding GILS elements. In addition, departmental database update procedures could complicate which records are selected for periodic conversion to GILS format. Further guidance and development of the conversion tool will be required to assist departments in managing their own data conversions.

Appendix E illustrates the mapping that was undertaken for a test conversion of InfoSource records to GILS.

4.2.3 Integrating GILS and Electronic Document Management

Based on industry projections and government statements of directions, a growing proportion of services and source information will be available in electronic formats. To accommodate this evolving electronic environment, a specification was developed for integrating GILS records in an electronic document management environment.

As demonstrated by the sample GILS records that were automatically generated for the Electronic Regulatory Filing initiative, the National Energy Board plus participating provincial agencies and industry sector companies will be producing GILS records automatically for SGML-based source documents. By leveraging an innovative construct, called architectural forms, these organizations will eliminate most of the manual labour and overhead associated with the creation of GILS records. Some professional effort will still be required to provide those elements that are not explicitly identified by SGML tags within source documents (e.g. controlled subject terms).

4.3 Evolution of Departmental GILS Sites

An underlying assumption of the target architecture was that it would evolve to include a growing number of departmental GILS databases as illustrated in Figure 2 below.

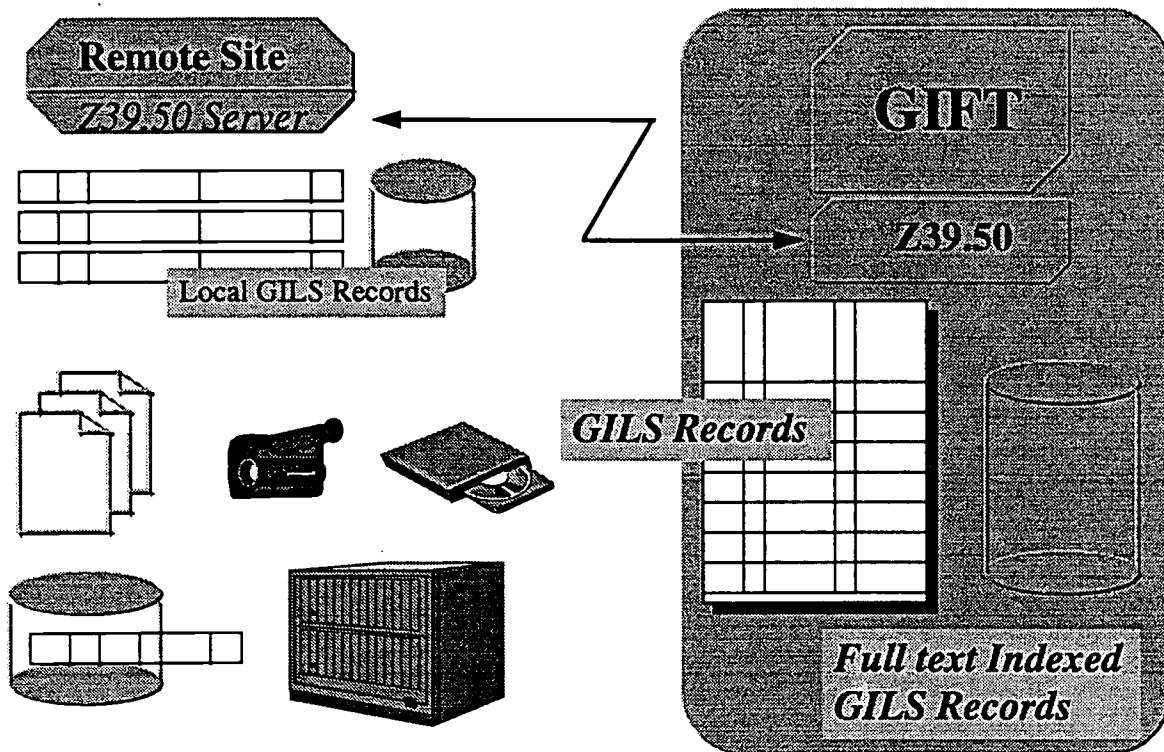


Figure 2 - Composite Departmental and Government-Wide GILS Environment

As illustrated above, the remote databases would contain GILS records as well as the information items themselves in a wide range of media types (e.g. CD-ROM, video, electronic and paper documents, etc.). The government's primary node on the Internet (denoted as GIFT) would hold a copy of GILS records maintained by each departmental database. It was further noted that the central registry might include only records for items that are publicly available. Users would get access to the departmental information holdings by identifying the location of the desired information before being switched, on request, to the departmental system to obtain a copy of the information item or to conduct a full text search of the item itself.

4.3.1 Information Services Enterprise (ISE)

In June 1996, TIMS approved a project to demonstrate inter-departmental co-operation in providing government information to users searching for a job or striving to develop an export market capability. Under the joint leadership of Industry Canada and Human Resources Development Canada, this initiative selected GILS as its standard for

describing information resources in these subject areas available to clients from a federation of government departments.

From a GILS perspective the project provided an enhanced user interface that enabled individuals to direct searches against one or more GILS databases. It gave users an ability to save the amalgamated search results at the local desktop for subsequent use. Preliminary indications are that creating a "federated" network of government information providers to identify, describe and provide information services in a common, standards-based manner is a manageable but challenging undertaking. These and other findings remain to be formalized and officially approved the project participants.

4.4 U.S. Implementation Experience

The second GILS Conference, held November 13-14, 1996 in College Park, Maryland, brought together over 300 delegates from across North America and Europe to discuss GILS implementations, issues and the future of GILS.

There was general consensus that GILS is a sound method for identifying and improving the dissemination government information and supporting information and record management. A GILS record is a "trusted" pointer to government resources because it not only describes the resource but also provides information about the creator of the record, date of record creation, originating department's programs and policies, contact name, etc.

There was also agreement that the technology for supporting GILS is not an issue. There are now more GILS compliant Z39.50 servers and clients than ever before. For example, the GILS system being developed for the Library of Congress will soon be placed in the public domain. This will include the Metadata Manager (administration tool), a Metadata Server (Z39.50/GILS server) and a Metadata Client (Z39.50 Java client). Even if a department does not have a Z39.50 server, it can make its GILS records available through Web technology.

Conference speakers revealed, however, that even though GILS is well entrenched in U.S. legislation (the Paper Reduction Act of 1995 states that department must provide a GILS service), there is great unevenness in the application of GILS by federal departments and agencies. Some have ignored GILS, while others have embraced it as the primary mechanism for identifying government resources. This unevenness is due primarily to the lack of a federal government-wide or even departmental-wide policy framework to support GILS. In addition, there are other competing initiatives and little co-ordination among the lead departments. (This situation is very similar to the one within the Canadian federal government.) The lack of a policy framework has resulted in mixed support by individual departments and a lack of consistent coverage by GILS records. Some departments have created less than 6 records, while other have created hundreds to identify all kinds of information resources. There were no guidelines on what should be identified by a GILS record.

There were several presentations on successful GILS implementations, including those by the U.S. Department of Defence, U. S. Environmental Protection Agency and Department of the Treasury. In all cases, success was achieved through corporate commitment, a policy framework and a team approach involving different sectors of the organization.

4.4.1 Implementation Factors

Following are some key factors that need to be addressed to ensure the future success of GILS:

4.4.1.1 Government and Department-wide Commitment to GILS

Legislation is not enough to ensure that GILS will be taken seriously. There must be a cultural change within a department to support public access to government information through GILS. Department management needs to establish GILS through policies and the assignment of the necessary resources to support a GILS service.

4.4.1.2 Department-wide Co-ordination

Different sectors within the organization responsible for information creation, management and distribution have to cooperate to support GILS. GILS needs to be team-driven through the participation of various sectors within the organization.

4.4.1.3 Content Guidelines

There must be government-wide guidelines indicating what information resources need to be identified. For example, GILS records should be created for all: Web sites, public assistance services, automated systems, publication catalogues, etc. The focus should be on what is important to the public. It is important to clearly state, on the GILS homepage, the scope of coverage of the described information resource, otherwise the public may be misled into believing that the GILS records identify everything that is available from the department.

4.4.1.4 Links to the Information Resource

It is not enough to just describe the information resource in a GILS record. To make a GILS record truly effective as an information service support tool, it should contain a hyperlink to the actual resource if that resource is available in an electronic format.

4.4.1.5 Information User Feedback

Government departments need to work with the public to determine what information is useful to government information users.

4.4.1.6 Integration of GILS into Web Home Pages

Include GILS into existing Web home pages and use GILS records whenever possible to identify government services and information.

4.4.1.7 Automated Generation of GILS Records

To reduce the effort required to create GILS record, record creation should be incorporated into the production of any electronic information resource.

4.4.1.8 Need to Market GILS

GILS needs to be better marketed to increase service demand. The government should be known for distributing information rather than restricting access.

4.4.1.9 Conference Conclusion

GILS is still in its infancy having been part of U.S. legislation for only eighteen months. The success of GILS, however, rests with corporate buy-in. A lot of work is now required by individual departments to put in place the policies and infrastructure to support GILS as a public service providing effective access to government information.

4.5 Other Governmental GILS Initiatives

In some circles GILS is starting to be referred to as the "global" rather than "government" information locator service. This may become the reality if other nations copy the efforts that have taken place in North America.

4.5.1 Australia

A recent Australian report, Architecture for Access to Government Information, report of the IMSC - Technical Group, recognizes that: "[GILS] is becoming a de facto International standard for the description of government information and this is likely to be of assistance to users of government information." and recommends use of GILS in the Australian government.

4.5.2 G7 Countries

The G7 Environment and Natural Resources Management (ENRM) Project has adopted Z39.50 with GILS and compatible profiles as the ENRM remote search standard, with GELOS-specific usage guidelines under development.

4.5.3 Other Initiatives

A full list of jurisdictions and international projects using or investigating GILS can be found at the U.S. Geological Survey site at:
<http://www.usgs.gov/public/gils/contacts.html>.

5. GILS Pilot Project - Phase I

5.1 Project Objectives

As established by the GILS Subgroup early into the GILS standard definition process, a pilot project was essential to:

- highlight the need for and the benefits of metadata;
- support user information discovery, and;
- assess the adequacy and utility of the GILS record as a metadata standard.

Phase I of the Pilot Project has focused on building a fairly small but representative database to assess the adequacy and utility of the GILS record and GILS profile (a Z39.50 customization to support user access to GILS data). In addition it provides a systems environment that helps the participants understand the environment and opportunities for enhancing access to departmental information resources through GILS records maintained in departmental databases and a central repository.

5.2 Project Participants

Participants in the Pilot Project were encouraged to create 30 or more records per department using the WEB-based application developed by GTIS (described in section 4.2.1). Alternatively the test contribution could be created using an in-house record creation facilities and the GTIS record conversion application (see section 4.2.2). Once the test database was in place the participants were provided with suggestions to exercise and evaluate the search facilities.

As of December 31, 1996, the participant departments with a firm commitment to perform the pilot tasks included:

- Agriculture and Agri-Food Canada;
- Canadian Heritage;
- Department of Finance;
- Environment Canada;
- Fisheries and Oceans;
- Indian and Northern Affairs;
- Industry Canada;
- National Archives of Canada;
- National Energy Board;
- National Library of Canada;
- National Research Council;
- Public Works and Government Services Canada including the Depository Services Program;
- Revenue Canada
- Statistics Canada, and;
- Treasury Board Secretariat.

5.3 Evaluation Exercise

The Project evaluation will assess the suitability of the GILS record and the effectiveness of the GTIS tools and identify needed improvements. The Pilot Project will determine the implications of creating, maintaining and updating GILS records. It will assist government departments to identify the skill sets required for record creation, estimate the resource requirements and place GILS in departmental workflows. Finally, the evaluation will attempt to assess the usefulness of GILS information to users of government information. For that, after all, is the *raison d'être* of GILS, to improve public access to government information.

The pilot was officially closed on February 14th and the evaluation and preliminary report writing started. All the outstanding tasks and the desired volumes of data could not be assembled and fully evaluated as anticipated. This task will continue through to the end of March. Assuming the availability of future funding the outstanding tasks will be rolled into the Phase 2 activities. In the meantime, the preliminary findings are given in the following sections.

All participants in the GILS Pilot Project were asked to complete questionnaires on record creation facilities and on search facilities. At time of writing this report, very few questionnaires have been returned. The following summaries are based on very small samples but they give an early indication of participant feedback.

5.4 Summary and Interpretation of Responses to Record Creation Facilities

Of the sixty-six record creators in the 1996-97 GILS Pilot Project, only nine responded to the GILS Record Creation Questionnaire. The following analysis is therefore an initial insight at best.

5.4.1 Profile of the Record Creators

Six of the nine respondents had prior knowledge of GILS in addition to that provided at the general information session given on October 30th. The other three had no prior knowledge of GILS. Four had cataloguing experience and five respondents indicated that cataloguing knowledge or experience would be an important skill for GILS record creators. This view was enforced in comments concerning the Canadian GILS Guidelines.

5.4.2 Record Creation Effort

Five respondents spent between 30 and 60 minutes creating a GILS record. All indicated that this time decreased as they gained experience. The only respondent without prior data management, cataloguing or data entry experience spent twice as much

time as the respondents with record creation experience. This respondent also created the most records for the GILS pilot.

Since six of these respondents had to consult outside sources to supply the mandatory elements, and still failed to complete them, it appears that information required for mandatory fields is not easily available to the record creator

5.4.3 Information Requirements

There was general agreement, that the mandatory elements were useful and, with the exception of "Availability" were fairly easy to create.

Among the optional elements, the spatial domain, bounding co-ordinates and sources of data were the least useful and most difficult to supply. Otherwise, the majority of optional elements were perceived as useful and fairly easy to create.

Although nearly all respondents felt that no additional mandatory fields are needed, suggestions were received to add: contents note for web pages; record type; and include "none" as a value for medium type.

5.4.4 Impact on Current Work Routines

All participants undertook GILS record creation as part of a special project. From the responses to date it is not possible to draw conclusions about the place of GILS in current work flows or its impact.

5.4.5 Utility of Record Creation Tool

All respondents used the html form provided by GTIS. Five indicated that this tool was difficult to use. Three reported initial problems that vanished in January as the tool was upgraded. Two respondents mentioned that the system had lost records that they had input.

While the system successfully returned the input records, in both SGML and HTML formats, the process was flawed because guidance was not provided on local storage and processing requirements. Thus the returned records were retained locally on diskettes, in e-mail folders, and on network files.

Record updating was tested by eight respondents. The majority found this task fairly easy to do and GTIS technical support satisfactory.

5.4.6 Adequacy of Canadian GILS Guidelines

Five of the respondents indicated that the Guidelines were difficult to use. They found them difficult to understand and work with. The language was too hard to understand, too cataloguing-like, and not created for the average lay person. The examples were the best feature.

5.4.7 Conclusions and Recommendations

The following preliminary recommendations can be drawn from these initial responses to the questionnaire:

- the Record Creation tool needs significant improvements;
- the GILS Guidelines need to be rewritten and oriented to average users;.
- the optional fields could be reviewed and possibly reduced in number, and;
- the ideal GILS record creator would be someone with prior experience in cataloguing.

Record creation is a complicated, time-consuming task made difficult by the insufficient information available on most departmental information resources. This is further complicated by the inadequacy of the record creation tool. Average government employees can not be requested to take 30 minutes to an hour to create a GILS record or expected to perform at this level of sophistication. It would be helpful to further automate the record-creation process when this is possible.

In short, the survey results have demonstrated that the GILS process needs simplification and that, in particular, the record creation tool and guidelines need improvement prior to any GILS implementation across the federal government.

5.5 Summary and Interpretation of Responses to Search Facilities

Registered users who had created GILS records were requested to complete the search survey. Response rate was 5.5%.

The searches performed were based on suggested search strategies provided by the GILS Subgroup. None of the respondents were novice searchers. Further evaluation of the search facility will be sought from government users that didn't participate in the pilot and from the public at large.

5.5.1 GILS Search Facility

Overall the respondents were pleased with the design of the search screens. Respondents found the speed of searching and the relevancy of the search results good.

5.5.2 Utility of GILS Data

The Mandatory Elements and the Abstract field were found to be the most useful fields in the Advanced Search.

5.5.3 Ancillary Needs and Features

The respondents found that the On-line Instructions were not simple, clear, nor consistent. The Help Information was not helpful nor consistent and could include more examples. It was badly written and contained spelling errors.

Problems occurred with the search mechanism itself (i.e. a term did not get included in the search result and in two instances it took the searcher to web sites and not to a GILS record).

6. Phase I Follow-Up and Downstream Development

During the past fifteen months, a significant number of issues and requirements have been identified. All substantive consideration was deferred to a subsequent phase. Given the will and conviction of the Subgroup that this project should proceed to the next phase, following are a list of additional items that will need to be addressed and resolved at a government-wide level.

6.1 Linkage With Other Information Delivery Services

Various information delivery services exist within government and new ones were initiated within the past year. The Depository Services Program publishes the Weekly Checklist of new government publications and includes links to Internet addresses. The National Library is a repository for all government publications and creates cataloguing records for these publications including location information in the AMICUS data base. InfoSource is the tool mandated under the Access to Information Act and the Privacy Act to describe the organization of the Government of Canada and its unpublished information holdings. Records in these three long-standing sources have been successfully mapped to GILS. A single window to government information will remain a vision, rather than become reality, until these independent systems are interfaced through the type of linkage, between systems and information items, provided by GILS.

Linkage of existing services with a legislative or policy mandate (InfoSource, National Library, DSP) with GILS will also provide a firm mandate for GILS and minimize duplication in data preparation by departments. Examples of this potential include the interface with Environment Canada and InfoSource described in section 4.2.2.1. The follow-up activities should investigate and demonstrate how other information systems and resources could be linked into a GILS network.

6.2 Thesaurus Support

Controlled subject terminology is an optional element in the GILS record and can be omitted to ease the record creation effort. However, as demonstrated in the ISE project, control subject terms are essential to effectively match information with user needs (e.g. for sector specific information). If this type of information is to be provided

with any degree of consistency, controlled lists of terms which can be used across government departments are essential to support information indexing and user information discovery.

6.3 Registration and Administration of Information Resource Identifiers

The Uniform Resource Locator or URL has become the defacto information resource identifier on the Internet. Since the URL is an electronic address a location within a device on which the information resource is stored, the URL will change whenever the item is relocated or removed from the system.

A formal system of resource identifiers must be instituted if government information resources are to be uniquely and persistently identified. Formal public identifiers (e.g. see ISO 9070) coupled with common entity reference management (e.g. see SGMLOpen catalogue) must be instituted to ensure that every information object is given a persistent and unique identifier. This system could be based on the standards-based approach being implemented by the National Energy Board and its energy sector partners or some other scheme. Allocation, registration and administration of such a system of formal public identifiers could become the responsibility of ISBN agencies or a GILS facility.

Accurate, unique and persistent identifiers will encourage users to move directly from the GILS record to the actual information resource and thereby promote self service and minimize the information service burden in departments.

6.4 Guidelines for Departmental Server Implementation

Deployment of GILS servers at departmental sites during the ISE pilot, clearly illustrated the need for technical expertise, co-ordination and planning. Various issues will arise which can best be resolved through a focal point that can schedule network installations, interoperability tests and is goal-oriented. This type of expertise and written guidelines will have to be available to support deployment of GILS servers across government.

6.5 Security and Authenticity of GILS Records

Users must be assured that they are being pointed to authentic government information. This is critical in situations with business, health, legal ramifications and obligations. Detailed guidelines for record creation, training , improved record creation tools (e.g. more detailed help, a spell checker. etc.) and departmental quality control procedures are essential to ensure high quality GILS records. In addition, as highlighted by the Communications Security Establishment appropriate provisions will need to be made to ensure that GILS records for certain resources will be authentic and secure against malicious or accidental change.

6.6 Enhancements to the Central Search Facility

The Pilot Project has identified numerous enhancements to the GILS search facility. In addition, a variety of features could be added to enhance the services and data maintained at the central GILS facility. Third party certification of information source or user authentication would facilitate the development of secure electronic commerce by supporting the interchange of legal documents. Central conversion services could reformat and restructure source information as required by the end users desktop applications. These examples illustrate the potential to make government information services more efficient and effective through inter-departmental collaboration and shared facilities.

The GILS comment button could ask the question: "Did you find the information you were looking for? If not, what were you seeking?" GILS should aim to meet the top information needs of the public.

6.7 Maintaining the Canadian GILS Guidelines

It is evident that additional examples and explanations are needed to assist record creators with understanding the purpose and structure of certain data elements. In addition, the guidelines will need to be extended as new elements and features are added to the GILS standard.

The maintenance could include interpretation support and advisory group consultation and resolution of ambiguous situations. A more effective link could be developed between the guidelines and on-line help for information discovery and record creation alike.

6.8 Development of Resource Description Guidelines

There is a real need to help information owners to better describe their information resources so that users will have a fairly uniform and somewhat consistent view of available government information and services. A resource description guideline, which recommended what information resources should be described and a what level of detail, would help government agencies deliver quality information services.

6.9 Additional Support for Archiving Function

Users such as the depository libraries have highlighted the need for reliable long term access to government electronic information resources. Archiving Internet information has been identified as an issue by the TB Internet Advisory Committee. The National Archives and the National Library are both committed to fulfil their respective mandates for preservation and access to electronic government records and publications.

A single element (i.e. schedule number) has been provided to identify an information resource that is archived according to federal government regulations.

Additional effort needs to be made to validate that this is sufficient or to define additional data and procedures that may be required.

6.10 Maintenance of Central GILS Database

It is clear that the technology exists to support a distributed GILS database configuration. However, the implications of moving to a distributed environment before GILS is operating smoothly need to be understood more fully before a decentralized option is made the preferred solution. The GILS target architecture, comprising a central database linked to departmental servers, provides a robust configuration that should provide reliable services during the initial development phases.

To ensure that government information remains accessible, procedures will be required to ensure that updated GILS records are retained at the central site whenever the information items or records are removed from departmental repositories.

6.11 Link to Directory Applications

A portion of the GILS record contains "contact" details and identifies the source where the actual information item can be obtained. This type of information appears in most government records and identifies the same "source" in many records. This type of information is also included in electronic directories. Investigations are underway to determine if electronic directory applications could communicate and interwork with GILS databases (i.e. interface the X.500 and Z39.50 protocols). If these efforts prove successful, GILS record creators will be provided with another opportunity to minimize the record creation effort by referencing the "contact" information held in a directory.

6.12 Communications Plan and Training Support

A communications plan and training program will need to be developed to support GILS throughout government.

The communication plan would advertise government strategies and services to make information more readily and easily available. It should focus on what is important to the public. GILS homepages and other media should clearly state the scope of coverage otherwise the public may be mislead into believing that the GILS records identify everything that is available from the department.

The training program on the otherhand would be directed at government departments and staff responsible for creating GILS records and maintaining the departmental information resources.

6.13 Integrating Cataloguing and GILS Records

Library cataloguing is rules based. The National Library and departmental libraries catalogue federal government publications according to international rules and using shared authority lists for names and subject terms. There are several options for integrating cataloguing with GILS thereby maximizing the quality of GILS records and minimizing duplication in record creation.

- Generate a GILS record from a cataloguing record.
- Library staff upgrade a GILS record created during the publishing process and convert the GILS record to a cataloguing record;
- A GILS record points to a library catalogue, not to individual publications, departmental publications are only described in the catalogue, and;
- An on-line library catalogue which is Z39.50 compliant could act as a GILS server.

Libraries must experiment with these options to determine which options are preferable under what conditions.

6.14 Integrating GILS with Information Workflow Management

As demonstrated by the National Energy Board and by Industry Canada, the capture of GILS descriptive information can be integrated into the document creation process. The extent and nature of this integration will be influenced by the departmental work environment and supporting technologies. At the most sophisticated level the GILS record can be extracted from the information identified in the document itself.

Alternatively, the GILS metadata can be captured as the document moves through its various stages of initial draft, revision and final release. Workflow management routines can use GILS metadata to identify and track the document throughout the entire information life cycle management. The combined workflow and GILS metadata could also be used to customize user views of the GILS record that is presented to the public and that is retained for internal use.

7. Recommendations

The GILS Pilot Project resulted in a number of successes but also highlighted important issues that need to be addressed to increase the probability that identification and location of government information resources will be effective, efficient and responsive to user needs. The future of GILS depends on the government-wide recognition and concerted action to ensure that all critical success factors are addressed on time and within available budgets. These critical factors can be categorized as policy, operational and technical. Strategies must be devised to ensure that each factor will be addressed and the required solutions will be implemented by individual departments and appropriate components included the government-wide information and information technology framework.

The following address specific aspects that need to be addressed in the policy, operation and technology spheres.

7.1 Policy Factors

7.1.1 Government-wide Commitment to GILS

Existing government policies need to be reviewed and adjusted as necessary or new ones developed to encourage government-wide implementation of GILS. These policies need to:

- establish and promote GILS as an official government standard;
- require GILS implementation as an official government policy;
- develop and implement a plan for the establishment of a centralized GILS facility and decentralized network of departmental GILS systems, and;
- establish a GILS Co-ordinator within each government department or agency.

7.1.2 Department-wide Co-ordination

In addition to government policy on GILS, a corporate culture must be fostered within departments to support government and public access to government information through GILS. Policies alone can not ensure that GILS will be taken seriously. Departmental management needs to establish GILS through internal policies and the assignment of the necessary personnel and infrastructure to support a GILS service. All sectors within a department responsible for information creation, management and distribution have to cooperate and adopt a team approach to achieve effective GILS implementation.

7.2 Operational Factors

The overall impact of a government-wide GILS infrastructure will be determined by the departmental commitment to create quality GILS records and to supplement this by providing effective access to the identified information resources. The added effectiveness and efficiency can be realized by:

7.2.1 Improving Information Access

Any ability to improve internal and external access will be heavily influenced by existing systems and information management practices. GILS implementation will improve information access if departments:

- implement a GILS server;
- provide automated links to the information resource;
- integrate GILS records with Web home pages;

- encourage user feedback, and;
- promote GILS as the preferred, common strategy for accessing government information.

7.2.2 Streamlining Record Creation Routines and Practices

Departments need to examine options and strive to achieve greater efficiencies in generating GILS records by:

- automated generation of GILS records;
- integrating descriptive data creation into electronic information workflow;
- reconciling and integrating library cataloguing activities with GILS record creation, and;
- integrating the information management routines through the use of common Intranet and Internet technologies.

7.2.3 Providing Staff Training and Development

Knowledgeable and committed employees are essential to improved services. The staff training and development at the department level would need to be co-ordinated and supported through government-wide training initiatives that would provide on-line help and training sessions. Introduction of GILS and the associated operational implications will require training and development in areas such as:

- implementing Z39.50 and setting up servers, gateways and client support;
- maintaining departmental information as a component of the government-wide information resources;
- providing enhanced access through use of thesaurus and controlled subject terms, and;
- creating and maintaining GILS records in a departmental and networked environment.

7.3 Technical Factors

There is a clear need for strong technical leadership and direction in the initial stages of GILS implementation. This guidance would best be provided by a "lead" agency that would be capable of:

- enhancing and maintaining a GILS record creation and search tools to incorporate new elements, values, edits, better bilingual support;
- maintaining and developing conversion tools to permit existing departmental records to be converted to the GILS format;

- extending the pilot GILS implementation to links with controlled vocabularies and shared authority lists to improve subject access to GILS records;
- updating and maintaining the Canadian GILS Guidelines;
- maintaining and developing the central GILS database to operate in a decentralized database environment, and;
- developing record security and authentication features in accordance with evolving government-wide standards and practices.

Appendix A - Members of the GILS subgroup

<u>Name</u>	<u>Dept.</u>	<u>E-Mail Address</u>
Cathy Blanchard	TBS	cathy.blanchard@tbs-sct.gc.ca
Wayne Bower	TBS	
Nancy Brodie	NLC Secretary	nancy.brodie@nlc-bnc.ca
Ed Buchinski	TBS Chair	buchinski.ed@tbs-sct.gc.ca
Pia Cole	EC	colep@cpgsv1.doe.ca
Diane Crouse	DFAIT	crouse.diane@extott06.x400.gc.ca
Sharron Curley	PWGSC	curleys@piper.pwgsc.gc.ca
Diana Dale	CH	diana_dale@pch.gc.ca
Tanis Dennis	DSP	dennist@dsp-psd.pwgsc.gc.ca
Ann Desormeaux	DFAIT	ann.desormeaux@extott06.x400.gc.ca
Nora Fontaine	DFAIT	nora.fontaine@extott06.x400.gc.ca
Bruno Gnassi	DSP	gnassib@dsp-psd.pwgsc.gc.ca
Oliver Javanpour	IC	javanpour.oliver@ic.gc.ca
Sharon Jeffrey	IC	jeffrey.sharon@ic.gc.ca
Fay McMullen	NAC	fmcullen@archives.ca
Tom Meyers	PWGSC	X.400
Dave Milne	DND	damilne@cals.debbs.ndhq.dnd.ca
Micheline Morissette	PWGSC	X.400
Debbie Roy-Pelletier	PWGSC	X.400
Ti Truesdale	DND	aa466@issc.debbs.ndhq.dnd.ca
Fay Turner	NLC	fay.turner@nlc-bnc.ca
Stephen White	GTIS	stephen.white@gta-atg.x400.gc.ca

Appendix B - Questionnaire for Record Creators

B.1 Profile of GILS record creator

Please indicate if you are completing this questionnaire for the GILS Subgroup Project _____ or ISE Working Group _____ or Both _____

1. Name and Department _____

2. Did you have any relevant field-creation experience Yes _____ No _____ (please check off as many as apply from the list below)

Cataloguing: _____
Data or Database Management: _____
Document Management: _____
Other, please specify _____

3. Did you have any prior knowledge of GILS. (Excluding general information session of Oct. 30th)

yes _____
no _____
some _____

4. Now that you have created some GILS records, what skills do you think would be important for a GILS record creator to have?

B.2 Time Spent Creating Records

5a. How many GILS records have you created for this pilot? _____

5b. How many of these were records using predominantly mandatory fields? _____

5c. How many of these were records using more than just the mandatory fields? _____

6a. On average, how long did it take to create a GILS record?

Less than 30 minutes _____
30 - 60 minutes _____
More than 60 minutes _____
Not sure _____

6b. Did this time decrease as you went along? Yes _____ No _____

6c. If you answered yes to the above, please indicate what was the least amount of time it took for you to create a single GILS record.

Less than 30 minutes _____
30 - 60 minutes _____
More than 60 minutes _____

7. On average, how long did other procedures related to GILS record production (e.g. sending files to GTIS, record keeping, consultation with author etc.) take (average time per record)?

Less than 5 minutes _____
5 - 10 minutes _____
10 - 30 minutes _____
More than 30 minutes _____
Not sure _____

8. How did the total amount of time involved in creating GILS records compare with your initial expectations (i.e. for gathering information, analyzing , inputting, and reviewing)? Did it take you:

Longer than expected _____
Less time than expected _____
About the same amount of time as expected _____
Not sure _____

B.3 Availability of Information Required

9. Were you able to complete the MANDATORY data elements for the information holding described:

Always _____
Most of the time _____
Some of the time _____
Never _____
If not Always, please indicate why not _____

10. Did you need to consult outside sources in order to complete the mandatory data elements for the information holding described?

Yes _____
No _____

11. If you answered "Yes" to question 2 above, please indicate which of the following outside sources you consulted when creating GILS records (please choose all which apply):

An existing record describing the information holding _____
The originator of the information holding _____
One of the GILS Subcommittee Group members _____
Other (please specify) _____

12. Please indicate the usefulness and ease of creation of the GILS data elements on the following 2 scales of 1 to 5 with 1 being most useful and 5 being least useful and easy to create. Please put a line through any of the data elements that you had no occasion to create.

MANDATORY ELEMENTS	USEFULNESS					EASE OF CREATION					
	1	2	3	4	5		1	2	3	4	5
Title	1	2	3	4	5		1	2	3	4	5
Originator	1	2	3	4	5		1	2	3	4	5
Date of Publication	1	2	3	4	5		1	2	3	4	5
Language of Resource	1	2	3	4	5		1	2	3	4	5
Availability	1	2	3	4	5		1	2	3	4	5
Medium	1	2	3	4	5		1	2	3	4	5
Distributor	1	2	3	4	5		1	2	3	4	5
Order Process	1	2	3	4	5		1	2	3	4	5
Order Information	1	2	3	4	5		1	2	3	4	5
Record Source	1	2	3	4	5		1	2	3	4	5
Language of Record	1	2	3	4	5		1	2	3	4	5
Date of Last Modification	1	2	3	4	5		1	2	3	4	5

OPTIONAL ELEMENTS	USEFULNESS					EASE OF CREATION					
	1	2	3	4	5		1	2	3	4	5
Author	1	2	3	4	5		1	2	3	4	5
Date of Publication Structured	1	2	3	4	5		1	2	3	4	5
Place of Publication	1	2	3	4	5		1	2	3	4	5
Abstract	1	2	3	4	5		1	2	3	4	5
Controlled Subject Index	1	2	3	4	5		1	2	3	4	5
Subject Terms Controlled	1	2	3	4	5		1	2	3	4	5
Controlled Term	1	2	3	4	5		1	2	3	4	5
Subject Terms Uncontrolled	1	2	3	4	5		1	2	3	4	5
Uncontrolled Term	1	2	3	4	5		1	2	3	4	5
Spatial Domain	1	2	3	4	5		1	2	3	4	5
Bounding Coordinates	1	2	3	4	5		1	2	3	4	5
West Bounding Coordinate	1	2	3	4	5		1	2	3	4	5
East Bounding Coordinate	1	2	3	4	5		1	2	3	4	5

OPTIONAL ELEMENTS	USEFULNESS	EASE OF CREATION
North Bounding Coordinate	1 2 3 4 5	1 2 3 4 5
South Bounding Coordinate	1 2 3 4 5	1 2 3 4 5
Place	1 2 3 4 5	1 2 3 4 5
Place Keyword Thesaurus	1 2 3 4 5	1 2 3 4 5
Place Keyword	1 2 3 4 5	1 2 3 4 5
Time Period	1 2 3 4 5	1 2 3 4 5
Time Period--Textual	1 2 3 4 5	1 2 3 4 5
Time Period--Structured	1 2 3 4 5	1 2 3 4 5
Beginning Date	1 2 3 4 5	1 2 3 4 5
Ending Date	1 2 3 4 5	1 2 3 4 5
Resource Description	1 2 3 4 5	1 2 3 4 5
Cost	1 2 3 4 5	1 2 3 4 5
Cost Information	1 2 3 4 5	1 2 3 4 5
Technical Prerequisites	1 2 3 4 5	1 2 3 4 5
Available Time Period	1 2 3 4 5	1 2 3 4 5
Available Time Textual	1 2 3 4 5	1 2 3 4 5
Available Time Structured	1 2 3 4 5	1 2 3 4 5
Beginning Date	1 2 3 4 5	1 2 3 4 5
Ending Date	1 2 3 4 5	1 2 3 4 5
Available Linkage	1 2 3 4 5	1 2 3 4 5
Linkage Type	1 2 3 4 5	1 2 3 4 5
Linkage	1 2 3 4 5	1 2 3 4 5
Sources of Data	1 2 3 4 5	1 2 3 4 5
Methodology	1 2 3 4 5	1 2 3 4 5
Access Constraints	1 2 3 4 5	1 2 3 4 5
General Access Constraints	1 2 3 4 5	1 2 3 4 5
Originator Dissemination Control	1 2 3 4 5	1 2 3 4 5
Security Classification Control	1 2 3 4 5	1 2 3 4 5
Use Constraints	1 2 3 4 5	1 2 3 4 5
Point of Contact	1 2 3 4 5	1 2 3 4 5
Purpose	1 2 3 4 5	1 2 3 4 5

OPTIONAL ELEMENTS	USEFULNESS					EASE OF CREATION				
Agency Program	1	2	3	4	5	1	2	3	4	5
Cross Reference	1	2	3	4	5	1	2	3	4	5
Cross Reference Title	1	2	3	4	5	1	2	3	4	5
Cross Reference Relationship	1	2	3	4	5	1	2	3	4	5
Cross Reference Linkage	1	2	3	4	5	1	2	3	4	5
Linkage Type	1	2	3	4	5	1	2	3	4	5
Linkage	1	2	3	4	5	1	2	3	4	5
Schedule Number	1	2	3	4	5	1	2	3	4	5
Original Control Identifier	1	2	3	4	5	1	2	3	4	5
Record Review Date	1	2	3	4	5	1	2	3	4	5

13. Did you find that there were any data elements needed to describe your information resource which were missing from the Mandatory elements?

Yes _____ (specify)
 No _____

14. Did you create abstracts for the information holdings you described?

Yes _____
 No _____

15. If you answered "No" to question 14 above, why not? Please check as many as apply.

Lack of time _____
 Lack of expertise _____
 The title(s) described the content sufficiently _____
 Other (please specify) _____

16a. Did you add subject terms to your GILS records?

No _____
 Yes _____

16b. If you used a thesaurus or an authoritative list, please indicate which one _____

16c. Did you use place names or an authoritative list ?

Yes _____
 No _____

BEST COPY AVAILABLE

B.4 Work Routines

17. Does your department already maintain records describing information holdings?

Yes _____

No _____

If you answered "No" to this question, proceed to Question 21.

18. To which of the following resources does your department currently contribute records describing information holdings?

InfoSource	_____
Departmental Library Catalogue	_____
National Library Union Catalogue	_____
Departmental Publications Catalogue	_____
Departmental Web Site	_____
Other (please specify)	_____

19. Do these records describe (please select all which apply):

Individual articles, chapters, sections, or graphics	_____
Individual monographs or serials	_____
Collections of information (a database, microfiche holdings, maps)	_____
Services (help desk)	_____
Other (please specify)	_____

20. If you are already using meta-data, please circle the level of difficulty encountered in adapting this data for GILS. (1 is most amount of difficulty and 5 least amount of difficulty).

1 2 3 4 5

B.5 Record Creation Tool

21. Did you use the record creation tool (html form) provided by GTIS?

Yes _____

No _____ (Please specify what you used)

If you answered "No" to this question, proceed to the next section (The Guidelines).

22a. On a scale of 1 to 5 (with 1 as most difficult and 5 as least difficult), how difficult was this tool to use?

1 2 3 4 5

22b. If you circled 1 or 2, please indicate why.

22c. In which format did you receive the return of the records? _____

22d. What did you do with the output records received? _____

23. Would you say that the amount of technical support received from GTIS was:

Satisfactory _____

Unsatisfactory _____ (please indicate why)

24a. Did you attempt to update records? Yes _____ No _____

24b. If yes, on a scale of 1 to 5, how easy was it to locate and update records once they were created and sent to the central repository? (1 = most difficult, 5 = least difficult)

1 2 3 4 5

25. Are there any specific recommendations you would like to make in terms of the design of this tool?

B.6 The Canadian GILS Guidelines

26a. On a scale of 1 to 5, how difficult to use are the Canadian GILS Guidelines?

(1 = most difficult, 5 = easy to use)

1 2 3 4 5

26b. If you circled 1 or 2, please indicate why_____

27. Please indicate what you liked BEST about this tool:

Amount of detail (appropriate amount of detail) _____

The examples _____

The format in which the information was presented _____

The access points (table of contents) _____

The language used to describe and explain GILS _____

Other (please specify) _____

28. Please indicate what you liked LEAST about this tool:

Amount of detail (too much detail _____ not enough detail _____)

The examples _____

The format in which the information was presented _____

The access points (table of contents) _____

The language used to describe and explain GILS _____

Other (please specify) _____

29. Are there any specific recommendations you would like to make for the improvement of these Guidelines (e.g., guidelines provided for language, structured date)?

Appendix C - Questionnaire for Search Participants

C.1 GILS Search Facilities

On a scale of 1-5, 1 being the highest, rate the following:

a) Design of the search screen 1 2 3 4 5
Comments:

b) Online instructions 1 2 3 4 5
Comments:

c) Help information 1 2 3 4 5
Comments:

d) Search mechanism 1 2 3 4 5
Comments:

e) Speed of searching 1 2 3 4 5
Comments:

f) Display of search results 1 2 3 4 5
Comments:

g) Relevancy of search results 1 2 3 4 5
Comments:

h) Content of GILS record 1 2 3 4 5
Comments:

2. General Comments: _____

3. Did you use only the Simple Search? Yes _____ No _____
If yes, skip to question 7.

4. In the Advanced Search, indicate the usefulness of the following mandatory elements:
1 = most useful, 5 = least useful, NA = didn't use

ELEMENT	SCALE
Title	1 2 3 4 5 NA
Originator	1 2 3 4 5 NA
Date of Publication (textual)	1 2 3 4 5 NA
Language of Resource	1 2 3 4 5 NA

Availability	1 2 3 4 5 NA
Medium	1 2 3 4 5 NA
Distributor Sub-Elements	1 2 3 4 5 NA
Order Process	1 2 3 4 5 NA
Order Information	1 2 3 4 5 NA
Control Identifier	1 2 3 4 5 NA
Record Source	1 2 3 4 5 NA
Language of Record	1 2 3 4 5 NA
Date of Last Modification	1 2 3 4 5 NA

5. Did you use any optional elements during your searches? Yes No
If No, skip to question 7.

6. In the Advanced Search indicate the usefulness of the following optional elements:
1 = most useful, 5 = least useful, NA = didn't use

ELEMENT	SCALE
Author	1 2 3 4 5 NA
Place of Publication	1 2 3 4 5 NA
Abstract	1 2 3 4 5 NA
Controlled Subject Index	1 2 3 4 5 NA
Subject Thesaurus	1 2 3 4 5 NA
Subject Terms Controlled	1 2 3 4 5 NA
Subject Terms Uncontrolled	1 2 3 4 5 NA
Spatial Domain	1 2 3 4 5 NA
Place & Place Sub-Elements	1 2 3 4 5 NA
Time Period	1 2 3 4 5 NA
Resource Description	1 2 3 4 5 NA
Cost Information	1 2 3 4 5 NA
Technical Prerequisites	1 2 3 4 5 NA
Available Time Period	1 2 3 4 5 NA

Available Time Textual	1 2 3 4 5 NA
Available Time Structured	1 2 3 4 5 NA
Available Linkage	1 2 3 4 5 NA
Linkage Type	1 2 3 4 5 NA
Linkage	1 2 3 4 5 NA
Sources of Data	1 2 3 4 5 NA
Methodology	1 2 3 4 5 NA
Access Constraints & Sub-Elements	1 2 3 4 5 NA
Use Constraints	1 2 3 4 5 NA
Point of Contact	1 2 3 4 5 NA
Supplemental Information	1 2 3 4 5 NA
Purpose	1 2 3 4 5 NA
Agency Programme	1 2 3 4 5 NA
Cross Reference & Sub-Elements	1 2 3 4 5 NA
Schedule Number	1 2 3 4 5 NA
Original Control Identifier	1 2 3 4 5 NA
Record Review Date	1 2 3 4 5 NA

C.2 General Questions

7. How would you describe your proficiency in on-line searching, where 5 would indicate a novice searcher?

1 2 3 4 5

8. Did the content of the GILS records lead you to an appropriate document repository, collection, service, system, etc.?

Yes _____ No _____ Comments:

9. Were official language requirements properly addressed?

Yes _____ No _____ Comments:

10. Are there any additional features you would like to see added to the search mechanism?

General Comments:

Appendix D - Bibliography

Australia. Information Management Steering Committee. Technical Group. Architecture for Access to Government Information; Report of the IMSC - Technical Group. 1996 <URL:<http://www.adfa.oz.au/DOD/imsc/imsctg/imsctg1a.htm>>

Canada. Canadian GILS Guidelines. Prepared by the National Library of Canada and the Treasury Board of Canada Secretariat, GILS Subgroup (GSG). Hull: 1996. <URL:http://gils.gc.ca/gils/info_ea.html>

Canada. Government of Canada Internet Guide. Ottawa: Treasury Board Secretariat, 1996. <URL: <http://canada.gc.ca/programs/guide/index.html>>

Canada. Guide d'Internet du gouvernement du Canada. Ottawa: Secretariat du Conseil du Trésor, 1996 <URL: <http://canada.gc.ca/programs/guide/index.html>>

Canada. Les lignes directrices relatives au Service de localisation de l'information du gouvernement régissant la préparation d'enregistrements GILS. 1966 <URL:http://gils.gc.ca/gils/guide_f.html> .

Canada. Government Telecommunications and Informatics Services. Creating GILS Records in an SGML Environment. 1996. <URL:http://gils.gc.ca/gils/creatingg_e.html>

Canada. Government Telecommunications and Informatics Services. GILS SGML DTD. Hull: Government Telecommunications and Informatics Services, 1996. <URL:http://gils.gc.ca/gils/info_ea.html>

Canada. Government Telecommunications and Informatics Services. Government Information Finder Technology & The Single-Window-Infrastructure. Ottawa: GTIS, 1996.

Christian, Eliot J. "GILS; What is it? Where's it going?" In: D-Lib Magazine. December 1996. <URL: <http://www.dlib.org/dlib/december96/12christian.html>>

Christian, Eliot. Why a Global Information Locator? <URL:<http://www.g7.fed.us/enrm/flyer.html>>

IFLANET. Metadata Resources. <URL: <http://www.nlc-bnc.ca/ifla/II/metadata.htm>>

Miller, Eric. Dublin Core Element Set Crosswalk. (DRAFT) 1997. <URL:<http://www.oclc.org:5046/~emiller/DC/crosswalk.html>>

Moen, William and Christian, Eliot. Application Profile For The Government InformationLocator Service (GILS). Version 2. 1996. <URL:http://www.usgs.gov/gils/prof_v2.html>

Sperberg-McQueen, C.M. and Lou Burnard. Guidelines for Electronic Text Encoding and Interchange (P3). 1994. <URL: <http://etext.virginia.edu/TEI.html> >

Turner, Fay. Z39.50 Information Retrieval Standard: Overview and Implementations. Ottawa: National Library of Canada, 1995 (Network Notes # 8) <URL:<http://www.nlc-bnc.ca/pubs/netnotes/notes8.htm> >

Turner, Fay. La norme Z39.50 de repérage de l'information : vue d'ensemble et application. Ottawa: Bibliothèque nationale du Canada, 1995 (Flash Réseau, Numéro 8) <URL: <http://www.nlc-bnc.ca/pubs/netnotes/fnotes8.htm> >

U.S. Library of Congress. Z39.50 Maintenance Agency Web page. <URL:<http://lcweb.loc.gov/z3950/agency/> >

Nancy Brodie, Fay Turner. "Government Information Locator Service (GILS) Developments in the Government of Canada" In: Liaison December 2, 1996

Nancy Brodie, Fay Turner "GILS: an Evolving International Approach to Enhancing Resource Discovery". RIAO'97, Montreal, June 25-27, 1997.

Appendix E - Mapping GILS and InfoSource Data Elements

The following table identifies the five different kinds of records maintained in InfoSource. It identifies the data elements in each of these records and gives the corresponding element in the GILS record. A blank entry in any cell indicates that a GILS element has no corresponding data element in InfoSource.

One or more codes appear, within brackets, after the GILS element name to indicate that an element is mandatory (i.e. M) or optional (i.e. O) and whether the element is repeatable (i.e. R) or not repeatable (i.e. NR)

<i>InfoSource</i>	<i>InfoSource</i>	<i>InfoSource</i>	<i>InfoSource</i>	<i>InfoSource</i>	<i>Canadian GILS</i>
<i>General Information</i>	<i>Program Records</i>	<i>Personal Information Banks</i>	<i>Manuals</i>	<i>Directory of Federal Govt Databases</i>	
Name of Department or Agency	Title	Bank Name	Title	Name of Database	TITLE (M, NR)
Name of Department or Agency	Name Organization	Name Organization	Organization	Originating Department	ORIGINATOR (M, R)
na					AUTHOR (O, R)
none	n.a. (when first registered - but not in record)	n.a.	not available	Date created	DATE OF PUBLICATION (M for publications or resources with discrete creation or update date, NR)
City from address of availability	City from address of availability	City from address of availability	City from address of availability	City from address of availability	PLACE OF PUBLICATION (O, NR)
English, French	English or French	English or French	English and French (unless labelled English only)	Language	LANGUAGE OF RESOURCE (M if applicable, R)
Background Responsibilities Organization	Description Access Note	Description Class Purpose Uses Notes	no	Type	ABSTRACT (O, NR)
					CONTROLLED SUBJECT INDEX (O, R)
	Topics			Subject coverage	SUBJECT TERMS UNCONTROLLED (O, NR)

BEST COPY AVAILABLE

<i>InfoSource</i>	<i>InfoSource</i>	<i>InfoSource</i>	<i>InfoSource</i>	<i>InfoSource</i>	<i>Canadian GILS</i>
<i>General Information</i>	<i>Program Records</i>	<i>Personal Information Banks</i>	<i>Manuals</i>	<i>Directory of Federal Govt Databases</i>	
na					SPATIAL DOMAIN (O, NR)
na				Geographic coverage	PLACE (O, R):
na					TIME PERIOD (O, R)
					AVAILABILITY (M, R):
none	Storage Medium (but free text)	default paper		Medium	MEDIUM (M, NR)
Departmental Public Enquiry point	ATIP Co-ordinator	ATIP Co-ordinator	ATIP Co-ordinator and Reading Rooms	Distributor	DISTRIBUTOR (M, NR):
na	PR Number	TBS Number Bank Number			RESOURCE DESCRIPTION (O, R)
Public Information Office (Source: Additional Information)	How to Apply (ATI)	How to Apply (Privacy)			ORDER PROCESS (M, NR)
			Reading Room		ORDER INFORMATION (M, NR)
No	Yes	Yes			COST (O, NR)
					COST INFORMATION (O, NR):
na					TECHNICAL PREREQUISITES (O, NR)
na		Retention			AVAILABLE TIME PERIOD (O, R)
na					AVAILABLE LINKAGE (O, R)
					LINKAGE TYPE (O, NR)
					LINKAGE (O, R)
na					SOURCES OF DATA (O, NR)
na					METHODOLOGY (O, NR)

<i>InfoSource</i>	<i>InfoSource</i>	<i>InfoSource</i>	<i>InfoSource</i>	<i>InfoSource</i>	<i>Canadian GILS</i>
<i>General Information</i>	<i>Program Records</i>	<i>Personal Information Banks</i>	<i>Manuals</i>	<i>Directory of Federal Govt Databases</i>	
na	Access to Info Exemptions & Exclusions	Privacy Act (p. VII) some personal information is confidential		Access	ACCESS CONSTRAINTS (O, NR)
na	Crown	Crown			USE CONSTRAINTS (O, NR)
na		Contacts		Contacts	POINT OF CONTACT (O, NR)
Legislation	Legislation	Legislation	Legislation		PURPOSE (O, NR)
na	Organization	Organization			AGENCY PROGRAM (O, NR)
na		Related to			CROSS REFERENCE (O, R)
		Program Record Number			CROSS REFERENCE TITLE (O, NR):
na		PAC Number			SCHEDULE NUMBER (O, NR)
created on conversion	Assign	Assign	Assign		CONTROL IDENTIFIER (M, NR)
					ORIGINAL CONTROL IDENTIFIER (O, NR)
TBS	InfoSource	InfoSource	InfoSource		RECORD SOURCE (M, NR)
English or French					LANGUAGE OF RECORD (M, NR)
Date of record creation	December 1996	December 1996	December 1996		DATE OF LAST MODIFICATION (M, NR)
					RECORD REVIEW DATE (O, NR)

BEST COPY AVAILABLE



REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: An Evaluation of the Federal Government's Implementation of the Government Information Locator Service (GILS) - Final Report

Author(s):

Corporate Source:

Publication Date:

June 30, 1997

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following two options and sign at the bottom of the page.

The sample sticker shown below will be
affixed to all Level 1 documents

Check here

For Level 1 Release:
Permitting reproduction in
microfiche (4" x 6" film) or
other ERIC archival media
(e.g., electronic or optical)
and paper copy.

The sample sticker shown below will be
affixed to all Level 2 documents

Check here

For Level 2 Release:
Permitting reproduction in
microfiche (4" x 6" film) or
other ERIC archival media
(e.g., electronic or optical),
but *not* in paper copy.

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL
HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS
MATERIAL IN OTHER THAN PAPER
COPY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

Level 1

Level 2

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

**Sign
here→
please**

Signature:

Printed Name/Position/Title:

Charles R. McClure, Ph.D.

Organization/Address:

Syracuse University
School of Information Studies
4-206 Center for Science & Technology
Syracuse, NY 13244-4100

Telephone:

(315) 443-2911

FAX:

E-Mail Address:

Date:

July 24, 1997

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:

.....
Address:

.....
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:

.....
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

**ERIC / IT
Center For Science & Technology
Room 4-194
Syracuse University
Syracuse, NY 13244-4100**

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to: